

GenCore version 5.1.6  
(c) 1993 - 2004 Compu

(without alignments)  
11505.225 million cell

..... 3082

ameters: 6747726

redicted by chance to have a score of the result being printed  
ical score distribution.

Accession	Gene	Species	Library	Insert Size (bp)	Read Length (bp)	Sequencing Method	Notes
AB054600	CDNA	enccod					
AB192087	Human	Tum					
ABx72012	DNA	enccod					
ABk71563	Human	diti					
AB192076	Human	Tum					
ABx72001	DNA	enccod					
ABgs4740	Genomic	D					
AAz36913	DNA	enccod					
AAsg0571	DNA	enccod					
AB192134	Mouse	Tum					
ABx72059	DNA	enccod					
ACH24215	Human	foe					
ACd39376	Human	col					
ACH15529	Human	adu					
ACG42325	Human	MAp					
AAz23032	Human	Kd3					
AAa49177	DNA	enccod					
AAz36914	DNA	enccod					
AAz36893	CDNA	enccod					
AAz36897	CDNA	enccod					
AAz23024	Rat	Kd312					
ADd34625	Mouse	mito					
AA124259	Probe	#14					

45	127.4	4.1	1108	8	ADA59035	Abd45035	Novel human
44	127.4	4.1	1108	4	AA159052	AA159052	Human pol.
43	127.4	4.1	1087	4	AA160838	AA160838	Human pol.
42	140.6	4.6	1249	6	ABO72825	ABO72825	Human
41	143.4	4.7	3061	6	ABO72848	ABO72848	Human
40	149.4	4.8	720	9	ADBO9579	ADBO9579	Human
39	151	4.9	3366	7	ABX34827	ABX34827	Human
38	151	4.9	1722	9	ADBO7335	ADBO7335	Novel
37	151	4.9	597	8	ACD26671	ACD26671	Human
36	165.8	5.4	1305	4	AB107789	AB107789	Human
35	188	6.1	405	7	ABX45108	ABX45108	Bovine
34	199.6	6.5	3079	2	AA233025	AA233025	Rat
33	234.6	7.6	3966	2	AA223023	AA223023	Human
32	283	9.2	368	6	ABSI17600	ABSI17600	Human
31	283	9.2	368	5	AA109818	AA109818	Probe
30	283	9.2	368	4	ABSA43113	ABSA43113	Human
29	283	9.2	368	4	AAK17661	AAK17661	Human
28	283	9.2	368	4	ABK43477	ABK43477	Human
27	283	9.2	368	4	ABA36323	ABA36323	Probe
26	283	9.2	368	4	ABA51382	ABA51382	Human
25	283	9.2	368	4	AA149541	AA149541	Human
24	283	9.2	368	4	ABA69388	ABA69388	Human

UTR 1	
54600	
AB554600 standard; cDNA; 3082 BP.	
AB554600;	
05-DEC-2002 (first entry)	
cDNA encoding human Ras-like protein.	
Human; Ras-like; cancer; apoptosis; cell proliferation; AIDS; allergy; acquired immunodeficiency syndrome; Alzheimer's disease; osteoporosis; Parkinson's disease; aplastic anaemia; atherosclerosis; inflammation; bacterial infection; viral infection; fungal infection; gene therapy; parasitic infection; protozoal infection; helminthic infection; Ras-inhibitor; Ras-stimulator; chromosome 22; gene; ss.	
Homo sapiens.	
Key	Location/Qualifiers
5'UTR	1..210
	/*tag= a
CDS	211..1011
	/*tag= b
	/product= "Human Ras-like protein"
3'UTR	1012..3082
	/*tag= c
WO200262849-A2.	
15-AUG-2002.	
28-JAN-2002; 2002WO-US002301.	
08-FEB-2001; 2001US-00778963.	
(PEKE ) PE CORP NY.	
Neelam B, Ketchum KA, Di Francesco V, Beasley EM;	
WPI; 2002-706901/76-	
P-PSDB; ABG70951.	
New human Ras-like polypeptides, useful for treating diseases associated with an increase in apoptosis or cell proliferation e.g., cancer.	

PS Claim 4; Fig 1A; 73pp; English.

The present invention relates to a new polypeptide comprising a fully defined 266-amino acid sequence or its fragment, orthologue or allelic variant. The allelic variant or orthologue is encoded by a nucleic acid molecule that hybridises under stringent conditions to the opposite strand of a nucleic acid molecule with a fully defined sequence of 1082 or 1121 bp. The pharmaceutical composition comprising the agent that binds to the polypeptide is useful for treating a disease or condition mediated by a human Ras-like protein e.g. cancer. Likewise, the polypeptide is useful for identifying modulators of Ras-like protein activity or expression by screening a compound for the ability to stimulate or inhibit interaction between the Ras-like protein and an interacting molecule or Ras-like activity. The polypeptide is useful for diagnosing a disease, or predisposition to a disease, or treating diseases associated with an increase in apoptosis or cell proliferation e.g. AIDS (acquired immunodeficiency syndrome), Alzheimer's disease, Parkinson's disease, aplastic anaemia, allergies, atherosclerosis, osteoporosis or bacterial, viral, fungal, parasitic, protozoal or helminthic infections, and inflammation. The present nucleic acid sequence encodes the human Ras-like protein of the invention. This gene is located on chromosome 22

Sequence 3082 BP; 708 A; 962 C; 793 G; 619 T; 0 U; 0 Other;

Query Match	100.0%;	Score 3082;	DB 6;	length 3082;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3082; Conservative	0;	Mismatches	0;	Gaps 0

OY	1	GGGCTGCGCGCGGGGAAAGAAAGACCGCGCCCAAGCCGGGTCTCCAGACAGCCCAAGG	60
Db	1	GGCGTCGCGCGCGGGGAAAGAAAGAACCGCGCCCAAGCCGGGTCTCCAGACAGCCGAAGG	60
OY	61	GAAGATCCCCCGGCGAGTGAACCGGGAGGCCACCAAGACTCTGGAGGCTCGGCGGTGGA	120
Db	61	GAAGATCCCCCGGCGAGTGAACCGGGAGGCCACCAAGACTCTGGAGGCTCTGGCGGTGGA	120
OY	121	GGAGCAGGAGCTCCCGCGAGCTCCCGGCGCTTCCAGGCAAGCTCTTGAGCGGTGCAGA	180
Db	121	GCAGCAGGAGAGCTCCCGCGAGCTCCCGGCGCTTCCAGGCAAGCTCTTGAGCGGTGCAGA	180
OY	181	GGCGCGGCGCGCGCATTTCCAGCGCCGAGCATAGATGAAGACTTTGTCCAGCGGGAACTGC	240
Db	181	GGCGCGGCGCGCGCATTTCCAGCGCCGAGCATATGAAGACTTTGTCCAGCGGGAACTGC	240
OY	241	ACGCTCAGTGTGCGCCGCCAAAATCTATACCGCATGGTGGTGTGGGTCTCTCGGGTG	300
Db	241	AGCTCAGTGTGCGCCGCCAAAATCTATACCGCATGGTGGTGTGGGTCTCTCGGGTG	300
OY	301	GGCAAGAGTCCATCGTGTCTGGCTTCTCTCAATGGCCGCTTTGAGGACCAAGTACAACC	360
Db	301	GGCAAGAGTCCATCGTGTCTGGCTTCTCTCAATGGCCGCTTTGAGGACCAAGTACAACC	360
OY	361	ACCATGAGGACTTCCACCGTAAGTATCAACATCCGCGGCGACATGTATCCAGCTCGAC	420
Db	361	ACCATGAGGACTTCCACCGTAAGTATCAACATCCGCGGCGACATGTATCCAGCTCGAC	420
OY	421	ATCTGGATACCTCTGTGGCAACACCCCTTCCCGCATGTGCGAGGCTGTCCACTTCACA	480
Db	421	ATCTGGATACCTCTGTGGCAACACCCCTTCCCGCATGTGCGAGGCTGTCCACTTCACA	480
OY	481	GGGAGTCTTCAATCCGTGGTTCAGCTGAGTATACCGGAGGCTTTCATGATGAGTCAAG	540
Db	481	GGGAGTCTTCAATCCGTGGTTCAGCTGAGTATACCGGAGGCTTTCATGATGAGTCAAG	540
OY	541	CGCCTTCAGAGCAGATCTTGAGAGTCAAGTCTCTGTAAGAACAAAGCAAGAGAGCG	600
Db	541	CGCCTTCAGAGCAGATCTTGAGAGTCAAGTCTCTGTAAGAACAAAGCAAGAGAGCG	600
OY	601	GGGAGCTGCCCATGTGTCATCTGTGGCAACAAGACGACCGGCAAGCTGTGCGCCGAG	660
Db	601	GGGAGCTGCCCATGTGTCATCTGTGGCAACAAGACGCGAGGCTGTGCGCCGAG	660

[illegible]

```

Db 1741 ATGGAATCATTTGTAAGTAAAGCTAGTACTCCCTTGAGCAGGCCCTACCCAG 1800
Qy 1801 TTTCAGATCAGGCGCTCCAGCCGGAGCGCTTCTCTCTGTGTCCCAACAGGGTTTCG 1860
Db 1801 TTTCAGATCAGGCGCTCCAGCCGGAGCGCTTCTCTCTGTGTCCCAACAGGGTTTCG 1860
Qy 1861 TGGCGTGTTCAGCTAGACATTTGACCTCCGCAATGAGCTCCACGGTTTACAGCAATT 1920
Db 1861 TGGCGTGTTCAGCTAGACATTTGACCTCCGCAATGAGCTCCACGGTTTACAGCAATT 1920
Qy 1921 GCACAGCGTGGGCTGGGCGCAGAGCTGCTTTTATATGCTCCCATTTTACAGAG 1980
Db 1921 GCACAGCGTGGGCTGGGCGCAGAGCTGCTTTTATATGCTCCCATTTTACAGAG 1980
Qy 1981 GATACACCGAGACTCGAGGGGACACGATAGACACAGGCCCATCTTGTCCCTTAGC 2040
Db 1981 GATACACCGAGACTCGAGGGGACACGATAGACACAGGCCCATCTTGTCCCTTAGC 2040
Qy 2041 AAATTCAGGGTACAGCTCCACTAGAACGAGCGCTGCTTACTGTGTCTGCTCTAG 2100
Db 2041 AAATTCAGGGTACAGCTCCACTAGAACGAGCGCTGCTTACTGTGTCTGCTCTAG 2100
Qy 2101 CATTATTTAAGCACCTACTGGGTGCTGGGTTCACTGTGTCTAGAGAAACAAAGGGTCC 2160
Db 2101 CATTATTTAAGCACCTACTGGGTGCTGGGTTCACTGTGTCTAGAGAAACAAAGGGTCC 2160
Qy 2161 CGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 CGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Qy 2221 GAGCGGGGAGAGACAGCTGGGAGCCCAAGATGAGCTGGGAGGATCCGACAGAAAGC 2280
Db 2221 GAGCGGGGAGAGACAGCTGGGAGCCCAAGATGAGCTGGGAGGATCCGACAGAAAGC 2280
Qy 2281 TCAGGGGCGGCTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2281 TCAGGGGCGGCTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Qy 2341 AGGCATGGTGGGTTTGAACCCATTAACCAAGGCGCTTGTATCAGCTCTTAACAGTA 2400
Db 2341 AGGCATGGTGGGTTTGAACCCATTAACCAAGGCGCTTGTATCAGCTCTTAACAGTA 2400
Qy 2401 TATTTGATTTTATCTCTCTTAAACATATGAAGTTTAAAGGCGCTTAAGAACCTTAGT 2460
Db 2401 TATTTGATTTTATCTCTCTTAAACATATGAAGTTTAAAGGCGCTTAAGAACCTTAGT 2460
Qy 2461 GATCTTCAATGGGCTCTTCTGAGGTTCAAGAGGGTAACTTCTCCAGGTCAAC 2520
Db 2461 GATCTTCAATGGGCTCTTCTGAGGTTCAAGAGGGTAACTTCTCCAGGTCAAC 2520
Qy 2521 AGCAAGTGTGGGTGGCAGAGCAAGCTAGCGCTGAGCATTCAGTACATACAGATGT 2580
Db 2521 AGCAAGTGTGGGTGGCAGAGCAAGCTAGCGCTGAGCATTCAGTACATACAGATGT 2580
Qy 2581 GCTCCCTCTCTTGAATGCTTGGGCGCTTGGGCGCTTGAAGGCTTGGGACATCTTGCTCA 2640
Db 2581 GCTCCCTCTCTTGAATGCTTGGGCGCTTGGGCGCTTGAAGGCTTGGGACATCTTGCTCA 2640
Qy 2641 ACCCTCTCCAGATCAGTCTGTGAGGGTCCCTGATGATTTGTTACACCATGCGCATG 2700
Db 2641 ACCCTCTCCAGATCAGTCTGTGAGGGTCCCTGATGATTTGTTACACCATGCGCATG 2700
Qy 2701 TATATACAGTACACAGATGTACACAGATGTACACAGATGTACATGCTCCAGGCTGCG 2760
Db 2701 TATATACAGTACACAGATGTACACAGATGTACACAGATGTACATGCTCCAGGCTGCG 2760
Qy 2761 ATACTGACCTGACCCCAAGCTTGGGCGCTTGGGCGCTTGGGCGCTTGGGCGCTTGGGCG 2820
Db 2761 ATACTGACCTGACCCCAAGCTTGGGCGCTTGGGCGCTTGGGCGCTTGGGCGCTTGGGCG 2820
Qy 2821 CCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Db 2821 CCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880

```

```

Db 2821 CCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Qy 2881 CTTTGGCTGCAACCGGTGTGTGGCCCGCTCTTACCCAGGACAGCCCGGCAATGATC 2940
Db 2881 CTTTGGCTGCAACCGGTGTGTGGCCCGCTCTTACCCAGGACAGCCCGGCAATGATC 2940
Qy 2941 TCCGTGTACTATCAATTAAGGTGGTTTGTACAAAGGTTTGTACAAAGGTTTGTACAA 3000
Db 2941 TCCGTGTACTATCAATTAAGGTGGTTTGTACAAAGGTTTGTACAAAGGTTTGTACAA 3000
Qy 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
Db 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
Qy 3061 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3082
Db 3061 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3082

RESULT 2
ABL92087
ID ABL92087 standard; cDNA; 2973 BP.
XX
AC ABL92087;
XX
DT 30-MAY-2002 (first entry)
XX
DE Human Tumour Endothelial Marker polynucleotide SEQ ID NO 197.
XX
KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostratic;
KW normal endothelial marker; pan-endothelial marker; immunostimulant;
KW antiangiogenic; tumour; neovascularisation; vascularised tumour;
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW psoriasis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200210217-A2.
XX
PD 07-FEB-2002.
XX
PF 01-AUG-2001; 2001WO-US024031.
XX
PR 02-AUG-2000; 2000US-0222599P.
PR 11-AUG-2000; 2000US-0224360P.
PR 11-APR-2001; 2001US-0282850P.
XX
PA (UWJO ) UNIV JOHNS HOPKINS.
XX
PI St Croix B, Kinzler KW, Vogelstein B;
XX
DR WPI; 2002-291856/33.
XX
P-PSDB; ABB90733.
XX
PT An isolated molecule comprising an antibody variable region which
PT specifically binds to an extracellular domain of a tumor endothelial
PT marker (TEM) protein, useful for inhibiting tumor growth.
XX
PS Claim 65; Page 152-153; 33pp; English.
XX
CC The invention relates to an isolated molecule comprising an antibody
CC variable region which specifically binds to an extracellular domain of a
CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
CC proteins have cytostratic, immunostimulant and antiangiogenic activity.
CC They are useful for inhibiting tumour growth, neovascularisation in subjects
CC bearing a vascularised tumour, polycystic kidney disease, diabetic
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)
CC are disclosed, as are marker oligonucleotide sequences: tumour
CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal
CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers
CC (PEM) ABL91903-ABL91995

```

XX Sequence 2973 BP, 596 A, 967 C, 787 G, 623 T, 0 U, 0 Other;  
SQ Query Match 95.7%; Score 2949.4; DB 6; Length 2973;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 GCCCGGCCAGCCCGGCTCCGAGACCGCAGGGAGAGATCCCGGCGCAGTGAACCCGGG 85  
DB 23 GCCCGGCCAGCCCGGCTCCGAGACCGCAGGGAGAGATCCCGGCGCAGTGAACCCGGG 82  
QY 86 AGCCACCAAGACTCTGGGAGGCTCGGCGCTGAGACAGAGGAGCTCCCGCAGCTCC 145  
DB 83 AGCCACCAAGACTCTGGGAGGCTCGGCGCTGAGACAGAGGAGCTCCCGCAGCTCC 142  
QY 146 CGGCGCTTCAAGCAGCTCTCTGAGCCGTCGACAGGCCCCGGCCGCTTCCAGCCCC 205  
DB 143 CGGCGCTTCAAGCAGCTCTCTGAGCCGTCGACAGGCCCCGGCCGCTTCCAGCCCC 202  
QY 206 GAGCCATGATGAAGCTTTGTCCAGCGGGAATGCAACCTCATGTGCCCCGAAAACT 265  
DB 203 GAGCCATGATGAAGCTTTGTCCAGCGGGAATGCAACCTCATGTGCCCCGAAAACT 262  
QY 266 CATACCGATGATGATGCTGGGCTGCTCGGCTGAGGCAAGAGCTCCATCGTCTCGCT 325  
DB 263 CATACCGATGATGATGCTGGGCTGCTCGGCTGAGGCAAGAGCTCCATCGTCTCGCT 322  
QY 326 TCTCTCAATGCGCGCTTTGAGAGCAAGTACACACCAACATCGAGGACTTCCACCTTAAG 385  
DB 323 TCTCTCAATGCGCGCTTTGAGAGCAAGTACACACCAACATCGAGGACTTCCACCTTAAG 382  
QY 386 TATACAACTCCGCGGCCCAATGTAACAGCTGGAATCCTGTGCAACCAACC 445  
DB 383 TATACAACTCCGCGGCCCAATGTAACAGCTGGAATCCTGTGCAACCAACC 442  
QY 446 CTTTCCCGGCGCAGGCTGTCATCTCAAGGGAGTCTTCACTCTGATGTTCA 505  
DB 443 CTTTCCCGGCGCAGGCTGTCATCTCAAGGGAGTCTTCACTCTGATGTTCA 502  
QY 506 GCTTGATTAACCGGAGTCTTCTGATGAGTCAAGGCTTCAAGAACGATCTGAGAG 565  
DB 503 GCTTGATTAACCGGAGTCTTCTGATGAGTCAAGGCTTCAAGAACGATCTGAGAG 562  
QY 566 TCAATCTCTGCTGAAGAACAGAACAGAGCGCGGAGCTGCCATGTCATCTGAG 625  
DB 563 TCAATCTCTGCTGAAGAACAGAACAGAGCGCGGAGCTGCCATGTCATCTGAG 622  
QY 626 GCAACAGAACGACACGCGGAGCTGTGCGCGCAGGTGCCCACACCGAGCGGAGCTGC 685  
DB 623 GCAACAGAACGACACGCGGAGCTGTGCGCGCAGGTGCCCACACCGAGCGGAGCTGC 682  
QY 686 TGTGTGCGGCGACGAGAACTCCGCTTCTTGAAGTGTGCGGCAAGAAACACCAAG 745  
DB 683 TGTGTGCGGCGACGAGAACTCCGCTTCTTGAAGTGTGCGGCAAGAAACACCAAG 742  
QY 746 TGAAGAGATGTTTCTAAGCTGCTTCAAGATGGCCAAAGTGCACACGAGATGAGCCCG 805  
DB 743 TGAAGAGATGTTTCTAAGCTGCTTCAAGATGGCCAAAGTGCACACGAGATGAGCCCG 802  
QY 806 CCTTGATCGCAAGATCTCCGTCAGTACGATGACGCTTCCACCCCAAGCCCTTCTGCA 865  
DB 803 CCTTGATCGCAAGATCTCCGTCAGTACGATGACGCTTCCACCCCAAGCCCTTCTGCA 862  
QY 866 TGGCGCGGCTCAAGAGATGAGAGCTTATGAGCATGTCTCGCCCTTGGCCGCCGCCCA 925  
DB 863 TGGCGCGGCTCAAGAGATGAGAGCTTATGAGCATGTCTCGCCCTTGGCCGCCGCCCA 922  
QY 926 GCGTCAACAGTACCTCAAGTACATCAAGGCCAAGGCTCTGCGGAGAGGCCAGGCCGCT 985  
DB 923 GCGTCAACAGTACCTCAAGTACATCAAGGCCAAGGCTCTGCGGAGAGGCCAGGCCGCT 982  
QY 986 AGAGGACAAGTGCACCATTCAGTGAAGGAGATGCTGAGGCGGAGCTTGGCAGATGCC 1045

DB 983 AGAGGACAAGTGCAACCATTCAGTGAAGGAGATGCTGAGGCGGCGCTTGGCCAGATGCC 1042  
QY 1046 TTGAGGAGAGTGGGCCCAAGATGCCCACTGAGCATATCCCAACGAGGCCCGGAGCA 1105  
DB 1043 TTGAGGAGAGTGGGCCCAAGATGCCCACTGAGCATATCCCAACGAGGCCCGGAGCA 1102  
QY 1106 GTCTTGTTCACAGACTTTAAGGACACAGCTGAGAGGCCCGGCGCTGAGCTCCGACAT 1165  
DB 1103 GTCTTGTTCACAGACTTTAAGGACACAGCTGAGAGGCCCGGCGCTGAGCTCCGACAT 1162  
QY 1166 TGTGTCTCTTCTCAAGCTTCTCTGAGTCCGCTTGTCCACAGCTCTTGTGTTCAT 1225  
DB 1163 TGTGTCTCTTCTCAAGCTTCTCTGAGTCCGCTTGTCCACAGCTCTTGTGTTCAT 1222  
QY 1226 CTCTCTGTGGAGAGACATCTCTGAGCTCAAGATGAGGAGTGAAGATTCGCTCA 1285  
DB 1223 CTCTCTGTGGAGAGACATCTCTGAGCTCAAGATGAGGAGTGAAGATTCGCTCA 1282  
QY 1286 CTTCTCTCTCGTGGGTTGGAAGAAATGTTGATGACAGAGGAGTGAAGATTCGCTCA 1345  
DB 1283 CTTCTCTCTCGTGGGTTGGAAGAAATGTTGATGACAGAGGAGTGAAGATTCGCTCA 1342  
QY 1346 TATGAGGCTTCTGAGAACAGCTTCAGATGAAGAACACAGAGGCCAGATGAAG 1405  
DB 1343 TATGAGGCTTCTGAGAACAGCTTCAGATGAAGAACACAGAGGCCAGATGAAG 1402  
QY 1406 GTCTCTCTCTCTGAGATTAACCCAGCTTGTGAGGAGGAGCTGGAGAACTTCTC 1465  
DB 1403 GTCTCTCTCTCTGAGATTAACCCAGCTTGTGAGGAGGAGCTGGAGAACTTCTC 1462  
QY 1466 TCCAGCCCTGCAACTCTTACGCTCTGATGAGTGCCTCTGCAACCCCTCCACCCCA 1525  
DB 1463 TCCAGCCCTGCAACTCTTACGCTCTGATGAGTGCCTCTGCAACCCCTCCACCCCA 1522  
QY 1526 GCAACACACAAATTTGGCCCCCAGCTGCTGACATTTAGAGCAGTGAATCTGTGCTG 1585  
DB 1523 GCAACACACAAATTTGGCCCCCAGCTGCTGACATTTAGAGCAGTGAATCTGTGCTG 1582  
QY 1586 AAGGGGAGGAGGACCAACCTCTAGACCAAGCCCACTTGAACCAAGCCCACTCTCTG 1645  
DB 1583 AAGGGGAGGAGGACCAACCTCTAGACCAAGCCCACTTGAACCAAGCCCACTCTCTG 1642  
QY 1646 ACCGCTTCTCAAGCTCTCTCTGAGTCCCTGCGCCGACAGTGTGCTTGTGTGG 1705  
DB 1643 ACCGCTTCTCAAGCTCTCTCTGAGTCCCTGCGCCGACAGTGTGCTTGTGTGG 1702  
QY 1706 TTGACAGCTTTTGTGTGATGATGATGATGAAGAAATCAATCTACTGTAAAGC 1765  
DB 1703 TTGACAGCTTTTGTGTGATGATGATGATGAAGAAATCAATCTACTGTAAAGC 1762  
QY 1766 CTAGTGAATCCCTCTTGTGAGGAGGCTCACCAGTTCAAGTCCAGCCCTCCACCCGG 1822  
DB 1763 CTAGTGAATCCCTCTTGTGAGGAGGCTCACCAGTTCAAGTCCAGCCCTCCACCCGG 1822  
QY 1826 ACGCTTCTCTCTGCTCCCAACAGGGTTCGCTGCTGTTGACGTTAGACATTTGA 1885  
DB 1823 ACGCTTCTCTCTGCTCCCAACAGGGTTCGCTGCTGTTGACGTTAGACATTTGA 1882  
QY 1886 CTTCCGCAATGAGTCCACGCTTTACAGAAATTTGACCAAGCGTGGGAGGAGGCA 1945  
DB 1883 CTTCCGCAATGAGTCCACGCTTTACAGAAATTTGACCAAGCGTGGGAGGAGGCA 1942  
QY 1946 GGAATGCTTTTAAATGCTCCCATTTACAGAGATACCAAGGAGCTCGAGGGAGC 2005  
DB 1943 GGAATGCTTTTAAATGCTCCCATTTACAGAGATACCAAGGAGCTCGAGGGAGC 2002  
QY 2006 ACGATGAGCAGGAGCCCACTTGTGCTCTAGCAAAATTCAGGGTCAAGCTTCACTAG 2065  
DB 2003 ACGATGAGCAGGAGCCCACTTGTGCTCTAGCAAAATTCAGGGTCAAGCTTCACTAG 2062  
QY 2066 AACGAGGCTGCTTACTGCTGCTGCTTCAAGCAATTTATTAAGCACTTACGAGGCTC 2125  
DB 2063 AACGAGGCTGCTTACTGCTGCTGCTTCAAGCAATTTATTAAGCACTTACGAGGCTC 2122



QY 2126 TGGGTTCACTGTCCTAGGAACCAAGAGGGTCCCAAGTCCGCGCTCTGCGCGCCCT 2185  
 Db 2123 TGGGTTCACTGTCCTAGGAACCAAGAGGGTCCCAAGTCCGCGCTCTGCGCGCCCT 2182  
 QY 2186 GCTGCCCAACCACTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGCACTGGAGCC 2245  
 Db 2183 GCTGCCCAACCACTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGCACTGGAGCC 2242  
 QY 2246 CAGAACTAGGCTGGGAGGAGATCCGACAGAAAAGCTCAGGGCGGGTCTTCTCTTGCC 2305  
 Db 2243 CAGAACTAGGCTGGGAGGAGATCCGACAGAAAAGCTCAGGGCGGGTCTTCTCTTGCC 2302  
 QY 2306 CGGGATTGGGGCTATGCTGGGTACCAACATGTACTCAGGCAATGGTGGGTTTGAACCCATA 2365  
 Db 2303 CGGGATTGGGGCTATGCTGGGTACCAACATGTACTCAGGCAATGGTGGGTTTGAACCCATA 2362  
 QY 2366 AACCAAGGCGCTTGTCTATCAGCTCTTAACAAGTATTTGTATTTTAACTCTCTAAA 2425  
 Db 2363 AACCAAGGCGCTTGTCTATCAGCTCTTAACAAGTATTTGTATTTTAACTCTCTAAA 2422  
 QY 2426 CATATTGAATTTTAGGCGCTTAAGAACTTATGATCTTCTATTTGGGCTCTTCTGAGG 2485  
 Db 2423 CATATTGAATTTTAGGCGCTTAAGAACTTATGATCTTCTATTTGGGCTCTTCTGAGG 2482  
 QY 2486 TTTCAGAGAGGTTAGTAATCTTCTCCAGGTACACAGCAAGTCTGTGGGTGGCAGAAACA 2545  
 Db 2483 TTTCAGAGAGGTTAGTAATCTTCTCCAGGTACACAGCAAGTCTGTGGGTGGCAGAAACA 2542  
 QY 2546 AGCTAGCGCTGGGCACTTCACTACATACCAAGATGTCCTCCCTCTTATGATGCTGGCC 2605  
 Db 2543 AGCTAGCGCTGGGCACTTCACTACATACCAAGATGTCCTCCCTCTTATGATGCTGGCC 2602  
 QY 2606 TGGGGCTTCAAGGGCTTTGGGACATCTTGTCTCAACCCCTCTCCCTAGATCAGTCTGGA 2665  
 Db 2603 TGGGGCTTCAAGGGCTTTGGGACATCTTGTCTCAACCCCTCTCCCTAGATCAGTCTGGA 2662  
 QY 2666 GGGTCCCTGTATGATTTGTATGTAACCAATGCCATATATATACAACTACACAGATGTAAC 2725  
 Db 2663 GGGTCCCTGTATGATTTGTATGTAACCAATGCCATATATATACAACTACACAGATGTAAC 2722  
 QY 2726 ACACAGATGTAACATGTCCTCAGAGCCCACTCTGATACCTGACCCAGCCCTT 2785  
 Db 2723 ACACAGATGTAACATGTCCTCAGAGCCCACTCTGATACCTGACCCAGCCCTT 2782  
 QY 2786 GGGCCCTGCTGCGTGTGTGCTCAAGAGACAGCTCCAAACCTGCTGTCCCTTCC 2845  
 Db 2783 GGGCCCTGCTGCGTGTGTGCTCAAGAGACAGCTCCAAACCTGCTGTCCCTTCC 2842  
 QY 2846 CACCACTGCTGAGAGCTTCTGAGAGACAGATACCTTTGGTGTGACCGGTGTGGCC 2905  
 Db 2843 CACCACTGCTGAGAGCTTCTGAGAGACAGATACCTTTGGTGTGACCGGTGTGGCC 2902  
 QY 2906 GCTCTCACCAAGGACAGAGCCCGGACCATGATCTCGGTGTAACATATATAAGTG 2965  
 Db 2903 GCTCTCACCAAGGACAGAGCCCGGACCATGATCTCGGTGTAACATATATAAGTG 2962  
 QY 2966 GGTGGTTTACA 2976  
 Db 2963 GGTGGTTTACA 2973

KW Tumour endothelial marker; normal endothelial marker; PEM;  
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
 KW neovascularization; immune response; cyclostatic; antidiabetic; gene;  
 KW ophthalmological; antineoplastic; antiarthritic; antipsoriatic; ds.  
 OS Homo sapiens.  
 PN W0200283874-A2.  
 PD 24-OCT-2002.  
 PF 10-APR-2002; 2002WO-US008253.  
 PR 11-APR-2001; 2001US-0282850P.  
 PR 06-FEB-2002; 2002US-0354262P.  
 PA (UWJO ) UNIV JOHNS HOPKINS.  
 PI Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;  
 DR WPI; 2003-093016/08.  
 DR P-PSDB; ABUS4440.  
 XX  
 XX  
 PT New purified human transmembrane protein, designated as tumor endothelial  
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors, or  
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.  
 PS Disclosure; Page 154-155; 374pp; English.  
 CC The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumours as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neovascularization or tumour angiogenesis, for  
 CC inducing an immune response to tumour endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumours. The present  
 CC sequence represents a human TEM or NEM gene of the invention  
 XX  
 SQ Sequence 2973 BP; 596 A; 967 C; 787 G; 623 T; 0 U; 0 Other;  
 Query Match 95.7%; Score 2949.4; DB 7; Length 2973;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 26 GCGCGGCCAGCCCGCGCTCCGAGACAGCGAGAGATCCCGCGCACTGACCCGGG 85  
 Db 23 GCGCGGCCAGCCCGCGCTCCGAGACAGCGAGAGATCCCGCGCACTGACCCGGG 82  
 QY 86 AGCCACCAAGACCTCTGGGAGGCTCGCGGTGAGAGACAGGAGCTCCCGGAGCTCC 145  
 Db 83 AGCCACCAAGACCTCTGGGAGGCTCGCGGTGAGAGACAGGAGCTCCCGGAGCTCC 142  
 QY 146 CGGCGCTTCCAGGAGCTCTCTGAGCGGTGACAGAGCCCGGCGCATTTCCAGAGCCC 205  
 Db 143 CGGCGCTTCCAGGAGCTCTCTGAGCGGTGACAGAGCCCGGCGCATTTCCAGAGCCC 202  
 QY 206 GAGCCATGATGAAGACTTTGTGACGCGGAACTGACAGCTCAGTGTGCGCGCAAAAAGT 265  
 Db 203 GAGCCATGATGAAGACTTTGTGACGCGGAACTGACAGCTCAGTGTGCGCGCAAAAAGT 262  
 QY 266 CATACCGCATGATGATGCTGGGAGGCTCTGCGGAGGAGAGAGTCCATGCTGTGCT 325  
 Db 263 CATACCGCATGATGATGCTGGGAGGCTCTGCGGAGGAGAGAGTCCATGCTGTGCT 322  
 QY 326 TCTCTCAATGGCGGCTTTGAGAGACAGTACACACCCATGAGAGACTTCCACCGTAAG 385  
 Db 323 TCTCTCAATGGCGGCTTTGAGAGACAGTACACACCCATGAGAGACTTCCACCGTAAG 382

RESULT 3  
 ABX72012  
 ID ABX72012 standard; DNA; 2973 BP.  
 XX  
 XX ABX72012;  
 XX  
 XX 12-MAR-2003 (first entry)  
 XX  
 XX DNA encoding human tumour endothelial marker TEM 2.  
 XX  
 XX Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;

QY 386 TATACATATCCGGGCAATGATACAGCTGCATCTCTGATATCTTTGGCAACACC 445  
Db 383 TATACATATCCGGGCAATGATACAGCTGCATCTCTGATATCTTTGGCAACACC 442  
QY 446 CTTTCCCGGCAATGGGAGGCTGTCCATCTTCAAGGGGATGTCTTCAATCCGTGGTTC 505  
Db 443 CTTTCCCGGCAATGGGAGGCTGTCCATCTTCAAGGGGATGTCTTCAATCCGTGGTTC 502  
QY 506 GCCTGATTAACCGGAGTCTTTCATGAGTTCAGAGCCCTTCAAGAGAGATCTCGAAG 565  
Db 503 GCCTGATTAACCGGAGTCTTTCATGAGTTCAGAGCCCTTCAAGAGAGATCTCGAAG 562  
QY 566 TCAAGTCTGCTTGAAGAACAGAACAGAGAGGCGGAGCTGCCCATGTGATCTGTG 625  
Db 563 TCAAGTCTGCTTGAAGAACAGAACAGAGAGGCGGAGCTGCCCATGTGATCTGTG 622  
QY 626 GCAACAAAGACACACGCGGAGCTGTGCGGAGGCTCCACACCGAGGCGGAGCTGC 685  
Db 623 GCAACAAAGACACACGCGGAGCTGTGCGGAGGCTCCACACCGAGGCGGAGCTGC 682  
QY 686 TGTGTCCGGGCGAGAGAACTCCGCTTCTTCAAGTGTCCGCAAGAGAACACCAACG 745  
Db 683 TGTGTCCGGGCGAGAGAACTCCGCTTCTTCAAGTGTCCGCAAGAGAACACCAACG 742  
QY 746 TGGAGAGATGTTCTAAGTCTTTCAGATGAGCCAGAGCTGCACACAGATGAGCCCG 805  
Db 743 TGGAGAGATGTTCTAAGTCTTTCAGATGAGCCAGAGCTGCACACAGATGAGCCCG 802  
QY 806 CCTTCATTCGCAAGATCTCCGTCAGTACGAGTACGAGCCCTTCCACCCGAGCCCTTCTGA 865  
Db 803 CCTTCATTCGCAAGATCTCCGTCAGTACGAGTACGAGCCCTTCCACCCGAGCCCTTCTGA 862  
QY 866 TCGCGCGGTCAAGAGATGAGCGCTATGAGTGTCTTCGCTTTCGCGCGCGCCCA 925  
Db 863 TCGCGCGGTCAAGAGATGAGCGCTATGAGTGTCTTCGCTTTCGCGCGCGCCCA 922  
QY 926 GCGTCAACAGTACTCTCAAGTATCAAGAGCCAGAGTCTTTCGAGAGGCGGCGCGG 985  
Db 923 GCGTCAACAGTACTCTCAAGTATCAAGAGCCAGAGTCTTTCGAGAGGCGGCGCGG 982  
QY 986 AGAGGAGCAAGTGCACATCCAGTGCAGAGGAGTGTGGGCGGGGCTTGGACAGTGC 1045  
Db 983 AGAGGAGCAAGTGCACATCCAGTGCAGAGGAGTGTGGGCGGGGCTTGGACAGTGC 1042  
QY 1046 TTCAGGAGGTGGCCCAAGATCCCATGTGTGCATCTTCCACCGAGGCGCGGCAACA 1105  
Db 1043 TTCAGGAGGTGGCCCAAGATCCCATGTGTGCATCTTCCACCGAGGCGCGGCAACA 1102  
QY 1106 GTCTTGTTCACAGACTTTCAGAGCAGAGTGCAGAGGCGCGGCGGCTTCCGCAACT 1165  
Db 1103 GTCTTGTTCACAGACTTTCAGAGCAGAGTGCAGAGGCGCGGCGGCTTCCGCAACT 1162  
QY 1166 TCGTCTGCTTCTCACAGCTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCAAT 1225  
Db 1163 TCGTCTGCTTCTCACAGCTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCAAT 1222  
QY 1226 CTCTCTGTGGAGGACATCTTCTGAGCTTCAAGAGTTAGGACAGACTCAAGTTACA 1285  
Db 1223 CTCTCTGTGGAGGACATCTTCTGAGCTTCAAGAGTTAGGACAGACTCAAGTTACA 1282  
QY 1286 CTTTCTCTCTGGGGTTCGAGAAATGTTTATGCGAGAGGGGTGAGAGTTCTGTGCTCA 1345  
Db 1283 CTTTCTCTCTGGGGTTCGAGAAATGTTTATGCGAGAGGGGTGAGAGTTCTGTGCTCA 1342  
QY 1346 TATGAGGCTCTGAGGACAGCTCAGATGAAAGAGACAGAAAGGCGAGATGAGAAAG 1405  
Db 1343 TATGAGGCTCTGAGGACAGCTCAGATGAAAGAGACAGAAAGGCGAGATGAGAAAG 1402  
QY 1406 GTCTCTCTCTCTGGCATTAACCCAGCTTGTGTTGGGTGAGAGCTGGAGAACTTCTC 1465  
Db 1403 GTCTCTCTCTCTGGCATTAACCCAGCTTGTGTTGGGTGAGAGCTGGAGAACTTCTC 1462

QY 1466 TCCAGGCTTCGCACTCTTAAGGCTCTGGTTCAGGCTGCAGACCCCTTCCACCCCA 1525  
Db 1463 TCCAGGCTTCGCACTCTTAAGGCTCTGGTTCAGGCTGCAGACCCCTTCCACCCCA 1522  
QY 1526 GCACACACAAAGTTGGCCCCAGCTGCCTGACATTTAGGCGCAGTGTGTCTG 1585  
Db 1523 GCACACACAAAGTTGGCCCCAGCTGCCTGACATTTAGGCGCAGTGTGTCTG 1582  
QY 1586 AAGGAGGCTGGCCACACTCTTTCAGACACGCGCCACACTTAAGACAGCCCACTCTG 1645  
Db 1583 AAGGAGGCTGGCCACACTCTTTCAGACACGCGCCACACTTAAGACAGCCCACTCTG 1642  
QY 1646 ACCGCTTCTCAGCTCTCTCTCCTAAGTCCCTCCGCGCCAGATTTGCTGTGTG 1705  
Db 1643 ACCGCTTCTCAGCTCTCTCTCCTAAGTCCCTCCGCGCCAGATTTGCTGTGTG 1702  
QY 1706 TTGAGCTGTTTGTGTATGATATAGTATGAAATGAAATCATTTGATCTGTAAAGC 1765  
Db 1703 TTGAGCTGTTTGTGTATGATATAGTATGAAATGAAATCATTTGATCTGTAAAGC 1762  
QY 1766 CTAGTACTCCCTCTCTTGGCCAGGCTTCAACCCAGTTCAAGTTCACGCGCTTCCACCGGG 1825  
Db 1763 CTAGTACTCCCTCTCTTGGCCAGGCTTCAACCCAGTTCAAGTTCACGCGCTTCCACCGGG 1822  
QY 1826 ACGCTTCTCTGCTGCTCCCAAGAGGTTCCGTGGCGCTTTGACAGTATGATTA 1885  
Db 1823 ACGCTTCTCTGCTGCTCCCAAGAGGTTCCGTGGCGCTTTGACAGTATGATTA 1882  
QY 1886 CCTCCGCAATGAGCTCCACGCTTTTACAGCAATTTGACAGAGCTGGGGGAGGAGCA 1945  
Db 1883 CCTCCGCAATGAGCTCCACGCTTTTACAGCAATTTGACAGAGCTGGGGGAGGAGCA 1942  
QY 1946 GCACTGCTTTTTTAAATGCTCCCATTTTCAAGAGAGTACACCGAGTCTGGAGGGAGC 2005  
Db 1943 GCACTGCTTTTTTAAATGCTCCCATTTTCAAGAGAGTACACCGAGTCTGGAGGGAGC 2002  
QY 2006 ACGATGAGCACAGGCGCCCACTTTGTCCCTTACCAATTCAGGGTACAGCTCCACCTAG 2065  
Db 2003 ACGATGAGCACAGGCGCCCACTTTGTCCCTTACCAATTCAGGGTACAGCTCCACCTAG 2062  
QY 2066 AACAGGCTGCTCTTCACTGTGCTGTTCTCAAGATTTTAAAGCACTTCTGGGTGC 2125  
Db 2063 AACAGGCTGCTCTTCACTGTGCTGTTCTCAAGATTTTAAAGCACTTCTGGGTGC 2122  
QY 2126 TGGGTTCATCTGTCTTTCAGAGAAACCAAGAGGCTCCAGTCTTGGCGCTTCCGCGCCCT 2185  
Db 2123 TGGGTTCATCTGTCTTTCAGAGAAACCAAGAGGCTCCAGTCTTGGCGCTTCCGCGCCCT 2182  
QY 2186 GCTGCCCAACACTTTCGACACACAGCGGTGGGAGGCGGGAGAGAGAGCTGGAGCC 2245  
Db 2183 GCTGCCCAACACTTTCGACACACAGCGGTGGGAGGCGGGAGAGAGAGCTGGAGCC 2242  
QY 2246 CAGAACTAGGCTTGGAGAGATTCGACAGAAAGCTCAGGCGGGTCTTCTCTGTGCGC 2305  
Db 2243 CAGAACTAGGCTTGGAGAGATTCGACAGAAAGCTCAGGCGGGTCTTCTCTGTGCGC 2302  
QY 2306 CCGGATTTGGGCTTATGCTGGGATACACAGATGATCAGGAGATGTTGAAACCCATA 2365  
Db 2303 CCGGATTTGGGCTTATGCTGGGATACACAGATGATCAGGAGATGTTGAAACCCATA 2362  
QY 2366 AACCAAGGCTTGTATCAGCTCTTAAACAAGTATTTTATCTCTCTTAA 2425  
Db 2363 AACCAAGGCTTGTATCAGCTCTTAAACAAGTATTTTATCTCTCTTAA 2422  
QY 2426 CATATTTAAGTTTAAAGGCTTAAAGAACTTAAAGTATCTTATGAGTCTTGTGAGG 2485  
Db 2423 CATATTTAAGTTTAAAGGCTTAAAGAACTTAAAGTATCTTATGAGTCTTGTGAGG 2482  
QY 2486 TTCAGAGAGGATTAATCTTCCAGAGTACACAGAGAGCTGTGGGTGGAGAGAAACA 2545  
Db 2483 TTCAGAGAGGATTAATCTTCCAGAGTACACAGAGAGCTGTGGGTGGAGAGAAACA 2542  
QY 2546 AGCTAGCCTGGGCAATTCAGTACATACAGAGATGTCTCTCTTGTATGCTTGGCCCC 2605

```

Db 2543 AGTATGCGCTGGGCAATTCAGTACATACACAGATGTCCTCTTATGCTTGGCCCC 2602
Qy 2606 TGGGGCCCTTCAGGGGCTTTGGGACATCTTGTCTCAACCTCTCCCTAATAGTCTGTGA 2665
Db 2603 TGGGGCCCTTCAGGGGCTTTGGGACATCTTGTCTCAACCTCTCCCTAATAGTCTGTGA 2662
Qy 2666 GGGTCCCTGTAGATTTGTGTACACATGACCATGATATATCAAGTACACAGATGAC 2725
Db 2663 GGGTCCCTGTAGATTTGTGTACACATGACCATGATATATCAAGTACACAGATGAC 2722
Qy 2726 ACACAGATGATACATGCTTCAGCCCGCAGCTCTGCATACCTGCACCTGACCCAGCCTT 2785
Db 2723 ACACAGATGATACATGCTTCAGCCCGCAGCTCTGCATACCTGCACCTGACCCAGCCTT 2782
Qy 2786 GGGCCCTTCGCTGCTGTGTGTCTCAAGACAGACGCTTCACCTGTCTGTCCCTTCC 2845
Db 2783 GGGCCCTTCGCTGCTGTGTGTCTCAAGACAGACGCTTCACCTGTCTGTCCCTTCC 2842
Qy 2846 CACCACTGCTGAGCCTTCTGAGACAGACAGTACCTTGGCTGACCGGTGTGGCC 2905
Db 2843 CACCACTGCTGAGCCTTCTGAGACAGACAGTACCTTGGCTGACCGGTGTGGCC 2902
Qy 2906 GCTCTCACCCAGGACAGAGCCCGCCAGCCATGATCTCCGTGTACACTATCAATAAAGTG 2965
Db 2903 GCTCTCACCCAGGACAGAGCCCGCCAGCCATGATCTCCGTGTACACTATCAATAAAGTG 2962
Qy 2966 GGTGTGTACA 2976
Db 2963 GGTGTGTACA 2973

```

## RESULT 4

```

ID ABR71563 standard; cDNA: 3427 BP.
AC ABR71563;
XX 30-JUL-2002 (first entry)
DT
XX
DE Human dlthp polynucleotide #29.
XX
KM Human; dlthp, diagnostic and therapeutic polynucleotide; gene; ss; bone;
KM cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
KM inflammatory disorder; viral infection; bacterial infection; seizure;
KM fungal infection; parasitic infections; developmental disorder; breast;
KM endocrine disorder; metabolic disorder; neurological disorder; cervix;
KM gastrointestinal disorder; transport disorder; gene therapy; kidney;
KM adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
KM skin; testis; thymus.
XX
OS Homo sapiens.
PN W0200220754-A2.
XX
PD 14-MAR-2002.
PF 29-AUG-2001, 2001WO-US027127.
XX
PR 05-SEP-2000; 2000US-0229747P.
PR 05-SEP-2000; 2000US-0229748P.
PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.
PR 05-SEP-2000; 2000US-0230583P.
PR 06-SEP-2000; 2000US-0230505P.
PR 06-SEP-2000; 2000US-0230514P.
PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.
PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.
PR 06-SEP-2000; 2000US-0230595P.
PR 06-SEP-2000; 2000US-0230597P.

```

```

PR 06-SEP-2000; 2000US-0230598P.
PR 06-SEP-2000; 2000US-0230599P.
PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230685P.
PR 06-SEP-2000; 2000US-0230988P.
PR 07-SEP-2000; 2000US-0230951P.
PR 07-SEP-2000; 2000US-0231163P.
PR 07-SEP-2000; 2000US-0231167P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL,
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR,
PI Montlyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM,
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Delfo A,
PI Marwana R, Chen AJ, Chang SC, Au AP, Imman RR,
XX
DR WPI; 2002-383054/41.
DR P-PSDB; ABG59971.
XX
PT An isolated polynucleotide useful in diagnostics and therapeutics.
XX
PS Claim 1; Page 419-420; 686pp; English.
XX
CC The invention relates to human diagnostic and therapeutic (dlthp)
CC polynucleotides and their associated polypeptides (dlthp polypeptides).
CC The sequences of the invention are used in the treatment and diagnosis of
CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal
CC infections, parasitic infections, developmental disorders (e.g. anaemia,
CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
CC (e.g. myotonic dystrophy, catatonla, peripheral neuropathy). Sequences
CC ABR71535-ABR71809 represent human dlthp polynucleotides of the invention
XX
SQ Sequence 3427 BP; 682 A; 1145 C; 935 G; 665 T; 0 U; 0 Other;
XX
Query Match 95.3%; Score 2937.2; DB 6; Length 3427;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2961; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
Qy 26 GCCGCGCCAGCCCGGCTCCCGAGCAGCGAGGAGATCCCGGCAAGTGAACCGGG 85
Db 459 GCCGCGCCAGCCCGGCTCCCGAGCAGCGAGGAGATCCCGGCAAGTGAACCGGG 518
Qy 86 AGCCACCAAGACTCTGAGAGGCTCGGCGCTGAGAGCAGAGCAGCTCCCGAGCTCC 145
Db 519 AGCCACCAAGACTCTGAGAGGCTCGGCGCTGAGAGCAGAGCAGCTCCCGAGCTCC 578
Qy 146 CGGCGCTTCAGGAGAGTCTCTAGCCGTGCAAGAGCCCGGCGCATTTCCAGCCCC 205
Db 579 CGGCGCTTCAGGAGAGTCTCTAGCCGTGCAAGAGCCCGGCGCATTTCCAGCCCC 638
Qy 206 GAGCGATGATGAAGACTTTGTCAGCGGGAATCTCAGCTCAGTGTCCCGCAAAAAC 265
Db 639 GAGCGATGATGAAGACTTTGTCAGCGGGAATCTCAGCTCAGTGTCCCGCAAAAAC 698
Qy 266 CATACCGCATGATGATGCTGAGTCTCTCTGAGGAGGAGTCCATGCTGTCTGCT 325
Db 699 CATACCGCATGATGATGCTGAGTCTCTCTGAGGAGGAGTCCATGCTGTCTGCT 758
Qy 326 TCCCTCAATGGCCGCTTGAAGACCAAGACACACCAATCGAGGACTTCCACCGTAAG 385
Db 759 TCCCTCAATGGCCGCTTGAAGACCAAGACACACCAATCGAGGACTTCCACCGTAAG 818
Qy 386 TATACCAATCCGCGGAGACATGATACAGCTCGATCTCGATACCTTGGCAACACC 445

```

Db 819 TATACACATCGCGGCGACATGTAACAGCTCGACATCTCTGGATACCTCTGGCAACACC 878  
QY 446 CTTTCCCGCCATGCGCAGGCTGTTCATCTCAAGGGGATGTCTTCACTCTGTGTTC 505  
Db 879 CTTTCCCGCCATGCGCAGGCTGTTCATCTCAAGGGGATGTCTTCACTCTGTGTTC 938  
QY 506 GCCTGATTAACCGGAGTCTTCCATGAGGTCAAGCGCTTCAAGAGAGATCCGAGG 565  
Db 939 GCCTGATTAACCGGAGTCTTCCATGAGGTCAAGCGCTTCAAGAGAGATCCGAGG 998  
QY 566 TCAAGTCTGCTGCTGAAGAACAAAGCAAGAGCGCGGAGAGTGCCTCATCTGTG 625  
Db 999 TCAAGTCTGCTGCTGAAGAACAAAGCAAGAGCGCGGAGAGTGCCTCATCTGTG 1058  
QY 626 GCAACAGAGACGACGCGGAGCTGTGCGCCGACAGTGCCTCAACGAGCGAGCTGC 685  
Db 1059 GCAACAGAGACGACGCGGAGCTGTGCGCCGACAGTGCCTCAACGAGCGAGCTGC 1118  
QY 686 TGGTGTGCGGCGACGAGAACTCCGCTACTTCGAGGTGTCCGCGCAAGAGAACACAAAG 745  
Db 1119 TGGTGTGCGGCGACGAGAACTCCGCTACTTCGAGGTGTCCGCGCAAGAGAACACAAAG 1178  
QY 746 TGGACGAGATGTTCTACGCTCTTCAAGCATGGCCAAAGCTGCACACGAGATGAGCCCCG 805  
Db 1179 TGGACGAGATGTTCTACGCTCTTCAAGCATGGCCAAAGCTGCACACGAGATGAGCCCCG 1238  
QY 806 CCTGCAATGCAAGATCTCCGTGCAAGTACGAGCTTCAACCCAGGCTCTTCTGCA 865  
Db 1239 CCTGCAATGCAAGATCTCCGTGCAAGTACGAGCTTCAACCCAGGCTCTTCTGCA 1298  
QY 866 TGGCGCGGCTCAAGAGATGAGAGCGCTAATGAGGTGTCCGCTTCCGCGCGCGCCCA 925  
Db 1299 TGGCGCGGCTCAAGAGATGAGAGCGCTAATGAGGTGTCCGCTTCCGCGCGCGCCCA 1358  
QY 926 GCGTCAACAGTACCTCAAGTACATCAAGGCCAAGGTCTTGGAGAGGCCAGGCCGCTG 985  
Db 1359 GCGTCAACAGTACCTCAAGTACATCAAGGCCAAGGTCTTGGAGAGGCCAGGCCGCTG 1418  
QY 986 AGAGGACAGTGCACCATCCAGTGAAGGAGATGCTGGGGCGGGGCTTGGCCAGTGC 1045  
Db 1419 AGAGGACAGTGCACCATCCAGTGAAGGAGATGCTGGGGCGGGGCTTGGCCAGTGC 1478  
QY 1046 TTCAGGAGAGTGGCCCAAGATGCCAATGGGAGCATCTCCACAGAGGCCCGCGAGCA 1105  
Db 1479 TTCAGGAGAGTGGCCCAAGATGCCAATGGGAGCATCTCCACAGAGGCCCGCGAGCA 1538  
QY 1106 GTCTTGTTCACAGACTTGAAGCAGACTGAGAGGCCCGCGGCGCTGCGCAAT 1165  
Db 1539 GTCTTGTTCACAGACTTGAAGCAGACTGAGAGGCCCGCGGCGCTGCGCAAT 1598  
QY 1166 TGGTGTGCTTCTCACAGCTTTCAGATCCGCTTGTCCACAGCTCTTGGTGTTCAT 1225  
Db 1599 TGGTGTGCTTCTCACAGCTTTCAGATCCGCTTGTCCACAGCTCTTGGTGTTCAT 1658  
QY 1226 CTCCTGTGAGGAGACACATCTGACAGCTCAAGATTAAGAGAGACTCAAGTTACA 1285  
Db 1659 CTCCTGTGAGGAGACACATCTGACAGCTCAAGATTAAGAGAGACTCAAGTTACA 1718  
QY 1286 CCTTCTCTCTGAGGGTGAAGAAATGTTGATCCAGAGGGGTGAGATTGCTGCTCA 1345  
Db 1719 CCTTCTCTCTGAGGGTGAAGAAATGTTGATCCAGAGGGGTGAGATTGCTGCTCA 1778  
QY 1346 TATGAGCTCTCTGAGGACAGCTCAAGATTAAGAGAGAGAGAGAGAGAGAGAG 1405  
Db 1779 TATGAGCTCTCTGAGGACAGCTCAAGATTAAGAGAGAGAGAGAGAGAGAGAG 1838  
QY 1406 GTCTCTCTCTCTGAGGACAGCTCAAGATTAAGAGAGAGAGAGAGAGAGAGAG 1465  
Db 1839 GTCTCTCTCTCTGAGGACAGCTCAAGATTAAGAGAGAGAGAGAGAGAGAGAG 1898  
QY 1466 TCCAGAGCTGCAAGCTTTAAGGCTGATGAGCTGCTGAGAGAGAGAGAGAGAGAG 1525  
Db 1899 TCCAGAGCTGCAAGCTTTAAGGCTGATGAGCTGCTGAGAGAGAGAGAGAGAGAG 1958

QY 1526 GCAACACACAGATGGGCCCCAGCTGGCCCTGACATTAAGAGAGAGAGAGAGAGAG 1585  
Db 1959 GCAACACACAGATGGGCCCCAGCTGGCCCTGACATTAAGAGAGAGAGAGAGAGAG 2018  
QY 1586 AAGGGGCGTGGCCACACTCTAGACACAGCCACCACTTAAGACAGCCCACTCTG 1645  
Db 2019 AAGGGGCGTGGCCACACTCTAGACACAGCCACCACTTAAGACAGCCCACTCTG 2078  
QY 1646 ACCGCTTCTCAGCTCTCTCTAGAGTCCCTCCGCGGACAGTGTGCTTGTGTG 1705  
Db 2079 ACCGCTTCTCAGCTCTCTCTAGAGTCCCTCCGCGGACAGTGTGCTTGTGTG 2138  
QY 1706 TTTGAGCTGTTTTGTGATGATAGTAAAGAAATGAAATCATTTGATTAAGAG 1765  
Db 2139 TTTGAGCTGTTTTGTGATGATAGTAAAGAAATGAAATCATTTGATTAAGAG 2198  
QY 1766 CTAGTGACTCCCTCTTGGCCAGGCTCTCAACCACTTCAAGTCAAGCTTCAACCCGG 1825  
Db 2199 CTAGTGACTCCCTCTTGGCCAGGCTCTCAACCACTTCAAGTCAAGCTTCAACCCGG 2258  
QY 1826 ACGCTTCTCTCTGCTTCCCAACAGAGTTTCCGTGGCTGTTGAGCTAGACATTA 1885  
Db 2259 ACGCTTCTCTCTGCTTCCCAACAGAGTTTCCGTGGCTGTTGAGCTAGACATTA 2318  
QY 1886 CCTCGCATTAAGCTCAAGCTTCAAGATTAAGACAAAGCTGGGGTGGGAGCCCA 1945  
Db 2319 CCTCGCATTAAGCTCAAGCTTCAAGATTAAGACAAAGCTGGGGTGGGAGCCCA 2378  
QY 1946 GAGCTCTTTTTTTAATGCTCCCATTTCAAGAGATTAACCAAGTCTGAGGGGAC 2005  
Db 2379 GAGCTCTTTTTTTAATGCTCCCATTTCAAGAGATTAACCAAGTCTGAGGGGAC 2438  
QY 2006 ACGATGAGACCAAGGCCCACTTGTCCCTTACCAATTCAGGGTACAGCTCCACCTAG 2065  
Db 2439 ACGATGAGACCAAGGCCCACTTGTCCCTTACCAATTCAGGGTACAGCTCCACCTAG 2498  
QY 2066 AACCAAGCTGCTCTACTGCTGCTGCTCTCAAGATTTAATTAAGACCTTACTGGGTG 2125  
Db 2499 AACCAAGCTGCTCTACTGCTGCTGCTCTCAAGATTTAATTAAGACCTTACTGGGTG 2558  
QY 2126 TGGGTTCACTGTGCTTAAGAAACCAAGAGGCTCCAGTCTGCTGCTGCTGCTGCT 2185  
Db 2559 TGGGTTCACTGTGCTTAAGAAACCAAGAGGCTCCAGTCTGCTGCTGCTGCTGCT 2618  
QY 2186 GCTGCCCCACCACTTTCGACACAGCGGTGGGGAGCGGGAGAGAGAGAGAGAGAG 2245  
Db 2619 GCTGCCCCACCACTTTCGACACAGCGGTGGGGAGCGGGAGAGAGAGAGAGAGAG 2678  
QY 2246 CAGAACTGAGCTGGAGAGATCCGACAGAAAGCTCAGGGCGGGTCTTCT - CCTGTG 2304  
Db 2679 CAGAACTGAGCTGGAGAGATCCGACAGAAAGCTCAGGGCGGGTCTTCT - CCTGTG 2738  
QY 2305 CCGGATTTGGGCTATGCTGGGATCCAGATTAAGAGAGAGAGAGAGAGAGAGAG 2364  
Db 2739 CCGGATTTGGGCTATGCTGGGATCCAGATTAAGAGAGAGAGAGAGAGAGAGAG 2798  
QY 2365 AAACCAAGGCGCTTGTATGAGCTCTTAACAGATTAATTTGATTTAATCTCTCA 2424  
Db 2799 AAACCAAGGCGCTTGTATGAGCTCTTAACAGATTAATTTGATTTAATCTCTCA 2858  
QY 2425 ACATATTAAGTTTTAGGGCCCTTAAGAACTTATGATGATTTGATTTGATTTGAT 2484  
Db 2859 ACATATTAAGTTTTAGGGCCCTTAAGAACTTATGATGATTTGATTTGATTTGAT 2918  
QY 2485 GTTCAAGAGGATTAAGTAACTTCTCCAGG - TCAACAGAGAGAGAGAGAGAGAG 2543  
Db 2919 GTTCAAGAGGATTAAGTAACTTCTCCAGG - TCAACAGAGAGAGAGAGAGAGAG 2978  
QY 2544 CAAGCTAGAGCTGGCATTTAGATCAATACAGATGCTGCTCTCTTGAATGCTGGCC 2603  
Db 2979 CAAGCTAGAGCTGGCATTTAGATCAATACAGATGCTGCTCTCTTGAATGCTGGCC 3038

QY	2604	CCTGGGGGCTTCAGGGCTTTGGGACATCTTGCCCTCAACCCCTCCAGATACAGTCTGT	2663
Db	3039	CCTGGGGGCTTCAGGGCTTTGGGACATCTTGCCCTCAACCCCTCTCCAGATACAGTCTGT	3098
QY	2664	GAGGGTCCCTGTAGATATTGTGTACACCATGCCATGTATATACAAAGTACACAGATGT	2723
Db	3099	GAGGGTCCCTGTAGATATTGTGTACACCATGCCATGTATATACAAAGTACACAGATGT	3158
QY	2724	ACACACAGATGTACACATGCTCCAGCCCCAGCTCTGACATACCTGTGACCTGTGACCCCAAGCC	2783
Db	3159	ACACACAGATGTACACATGCTCCAGCCCCAGCTCTGACATACCTGTGACCTGTGACCCCAAGCC	3218
QY	2784	TTGGGCCCCCTGGCTGTGGTCTGTGCTTAAAGCAGACAGCTCCAAACCTGGCTCTGTCCCCCTTC	2843
Db	3219	TTGGGCCCCCTGGCTGTGGTCTGTGCTTAAAGCAGACAGCTCCAAACCTGGCTCTGTCCCCCTTC	3278
QY	2844	CCACACCACTGCTCAGGCTCTCTGAGCAGACCAAGTACCTTGGGCTGACCCGATGTTGGC	2903
Db	3279	CCACACCACTGCTCAGGCTCTCTGAGCAGACCAAGTACCTTGGGCTGACCCGATGTTGGC	3338
QY	2904	CCGCTCTCACCCAGGCAGGCCCCGCCACCAATGATCTCCGTGTACATATCAATATAAAG	2963
Db	3339	CCGCTCTCACCCAGGCAGGCCCCGCCACCAATGATCTCCGTGTACATATCAATATAAAG	3398
QY	2964	TGGGTTTGTTCACAAAAAATAAAAA 2989	
Db	3399	TGGGTTTGTTCACAAAAAATAAAAA 3424	

RESULT 5	
ABL92076	
ID	ABL92076 standard; cDNA; 2832 BP.

DT 30-MAY-2002 (first entry)

DE Human Tumour Endothelial Marker polynucleotide SEQ ID NO 174.

KM Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytoskeletal;  
 KM normal endothelial marker; pan-endothelial marker; immunostimulant;  
 KM antiangiogenic; tumour; neoangiogenesis; vascularised tumour;  
 KM polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;  
 KM psoriasis; gene; ss.

05 Homo sapiens.

PN WO200210217-A2

PD 07-FEB-2002.

PF 01-AUG-2001; 2001WO-US024031.

PR 02-AUG-2000; 2000US-0222599P.

PR 11-APR-2001; 2001US-028285OP.  
VY

PA (UYJO ) UNIV JOHNS HOPKINS.  
XY

Pl St Croix B, Kinzler KW, Vogelstein B,  
XX

WPL; 2002-291856/33.  
DR  
XX

PT specifically binds to an extracellular domain of a tumor endothelial

XX  
L E M A N T O C I P R O C C H I / G B E T U R L O T Y M M M M M M M M

DISCLOSED, PAGE 110-120, COPY, ENLIGHTEN.

variable region which specifically binds to an extracellular domain of a

CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM

CC proteins have cytoprotective, immunostimulant and antiangiogenic activity.  
CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects  
CC bearing a vascularised tumour, polycystic kidney disease, diabetic  
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM  
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)  
CC are disclosed, as are marker oligonucleotide sequences; tumour  
CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal  
CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers  
CX (PEM) ABL91903-ABL91995  
XX  
SQ Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 U; 0 Other;

Query Match	91.1%;	Score 2806.4;	DB 6;	Length 2832;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2807; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	199	AAGCCCCGAGCATTGATGAAGA	CTTTGTCCAGCGGGAAATCGCA	AGCTCAGTGTCCCGC	258
Db	25	CAGCCCCGAGCATTGATGAAGAT	CTTTGTCCAGCGGGAAATCGCA	AGCTCAGTGTCCCGC	84
OY	259	AAAACTCATACCGCATGTGTGT	GTCTGGTGTCCCTCTCGGTTGGCA	AAAGCTTCATGTGT	318
Db	85	AAAACTCATACCGCATGTGTGT	GTCTGGTGTCCCTCTCGGTTGGCA	AAAGCTTCATGTGT	144
OY	319	TCTGTCTTCTCATATGGCCG	TTTGAGGACAGTACACCCACCAT	TGAGAACTTCCAC	378
Db	145	TCTGTCTTCTCATATGGCCG	TTTGAGGACAGTACACCCACCAT	TGAGAACTTCCAC	204
OY	379	CGTAAGGTATCAACATCCGCG	GCGATGTACAGCTCGACATCT	CGATCTTGAC	438
Db	205	CGTAAGGTATCAACATCCGCG	GCGATGTACAGCTCGACATCT	CGATCTTGAC	264
OY	439	AACCACTTTCCTCCGCGCAT	GTGCGGAGCTGTCCATCTCA	CAGGGAGTGTCTTATCTGT	498
Db	265	AACCACTTTCCTCCGCGCAT	GTGCGGAGCTGTCCATCTCA	CAGGGAGTGTCTTATCTGT	324
OY	499	GTGTTCAGCCGTGATACCG	GGGAGTCTCTTGATAGGTCA	AGCGCTTCAGAACAGATC	558
Db	325	GTGTTCAGCCGTGATACCG	GGGAGTCTCTTGATAGGTCA	AGCGCTTCAGAACAGATC	384
OY	559	CTGAGGTCAGAGTCTTGCT	GATGAAACAAGCAAGGAGCG	GCGAGCTGCCATGTCT	618
Db	385	CTGAGGTCAGAGTCTTGCT	GATGAAACAAGCAAGGAGCG	GCGAGCTGCCATGTCT	444
OY	619	ATCTGTGGCAACAAGAAC	CAACACGCGGAGCTGTGTCC	CGCAGAGTGTCCCA	678
Db	445	ATCTGTGGCAACAAGAAC	CAACACGCGGAGCTGTGTCC	CGCAGAGTGTCCCA	504
OY	679	GAGCTGTGTGTGTGGGCG	AGCAAGCAATCCGCGCTACT	TCGAGTGTGTGGCGCA	738
Db	505	GAGCTGTGTGTGTGGGCG	AGCAAGCAATCCGCGCTACT	TCGAGTGTGTGGCGCA	564
OY	739	ACCAACGTGACGAGATG	TTCTACTAGTCTCTTCAGAC	TGGCCAACTGCCACAGATG	798
Db	565	ACCAACGTGACGAGATG	TTCTACTAGTCTCTTCAGAC	TGGCCAACTGCCACAGATG	624
OY	799	AGCCCCCGCCGTGATCC	CAAGATTCCTCGTGTAGTAC	CGCTTCACACCCAGCGCC	858
Db	625	AGCCCCCGCCGTGATCC	CAAGATTCCTCGTGTAGTAC	CGCTTCACACCCAGCGCC	684
OY	859	TTCTGCATATGCGCCG	CTCAAGAGATGAGCGCT	ATGTCTCGCCCTTTCGCGCG	918
Db	685	TTCTGCATATGCGCCG	CTCAAGAGATGAGCGCT	ATGTCTCGCCCTTTCGCGCG	744
OY	919	CGCCCCCAGCTCAAC	TGTAACCTCAAGGCA	AGTGTCTTTCGGGAAAGCCAG	978
Db	745	CGCCCCCAGCTCAAC	TGTAACCTCAAGGCA	AGTGTCTTTCGGGAAAGCCAG	804
OY	979	GCCCGTGAAGAGGACA	AGTGTACCATTCAGGAGCA	AGGATGTCTGGGCGGGGCTTGTGC	1038
Db	805	GCCCGTGAAGAGGACA	AGTGTACCATTCAGGAGCA	AGGATGTCTGGGCGGGGCTTGTGC	864
OY	1039	CAGTGTCTTACGAGAG	TGGCGCCCAAGTATGCCCAT	GTGTGCGCATCTCCCCAC	1098

```

Db      865  CAGTGCTTCAAGGAGGTGGCCAGATGCCAATGTGGCAATCTCCACCGAGGCCCC 924
Qy      1099  GCGAGACGCTTGTGTACAGACCTTAGGACCAAGACTGGAGGCCCGCGGTGGCTTC 1158
Db      925  GCGAGACGCTTGTGTACAGACCTTAGGACCAAGACTGGAGGCCCGCGGTGGCTTC 984
Qy      1159  GCGACATTTGTGCTCTCAAGCTTCTGAGTCCGCTTGTCCACAGCTCTTGGTGG 1218
Db      985  CCGACATTTGTGCTCTCAAGCTTCTGAGTCCGCTTGTCCACAGCTCTTGGTGG 1044
Qy      1219  GTTTCATCTCTCTGTGGAGGACACATCTGACGCTCAAGAGTTAGGACAGACTGA 1278
Db      1045  GTTTCATCTCTCTGTGGAGGACACATCTGACGCTCAAGAGTTAGGACAGACTGA 1104
Qy      1279  AGTTACACTTCTCTCTGTGGAGGAGAAATGTTATGCCAGAGGGGTGAGATTGC 1338
Db      1105  AGTTACACTTCTCTCTGTGGAGGAGAAATGTTATGCCAGAGGGGTGAGATTGC 1164
Qy      1339  TGCATATATGAGGCTCTGTGGAGCAAGCTCAGATGAAAGAGACAGAGGCCAGAT 1398
Db      1165  TGCATATATGAGGCTCTGTGGAGCAAGCTCAGATGAAAGAGACAGAGGCCAGAT 1224
Qy      1399  GAGAAAGTCTCTCTCTGTGGCAATACACCAAGCTTGTGGTGGAGCTGGAGA 1458
Db      1225  GAGAAAGTCTCTCTCTGTGGCAATACACCAAGCTTGTGGTGGAGCTGGAGA 1284
Qy      1459  ACTTCTCTCCAGGCTCTGCACTTACGCTGTGGTTCAGCTGCTGACACCTCTCC 1518
Db      1285  ACTTCTCTCCAGGCTCTGCACTTACGCTGTGGTTCAGCTGCTGACACCTCTCC 1344
Qy      1519  ACCCCAGACACACAGATTGAGCCCGAGCTGCGCTGACATTGAGCGATGACTCT 1578
Db      1345  ACCCCAGACACACAGATTGAGCCCGAGCTGCGCTGACATTGAGCGATGACTCT 1404
Qy      1579  GTGTGAGAGGGGGGTGGCCACACTCTAGACCAAGCCCAACTTACAGACGCCCC 1638
Db      1405  GTGTGAGAGGGGGGTGGCCACACTCTAGACCAAGCCCAACTTACAGACGCCCC 1464
Qy      1639  CCTCTGACCGGCTTCTGAGCTCTCTCTGAGTCCCTCCGCGGACAGTTGCTT 1698
Db      1465  CCTCTGACCGGCTTCTGAGCTCTCTCTGAGTCCCTCCGCGGACAGTTGCTT 1524
Qy      1699  GTTGTGATGAGCTGTTTCTGATGATAGTAGAAATGGAATCAATTGACTG 1758
Db      1525  GTTGTGATGAGCTGTTTCTGATGATAGTAGAAATGGAATCAATTGACTG 1584
Qy      1759  TAAAAAGCTAGTACTCCCTCTGTGGCCAGGCTTCAACCAAGTTCAAGCTCC 1818
Db      1585  TAAAAAGCTAGTACTCCCTCTGTGGCCAGGCTTCAACCAAGTTCAAGCTCC 1644
Qy      1819  ACCCGGAGCGCTTCTCTCTGTGGCCAGGCTTCAACCAAGTTCAAGCTCC 1878
Db      1645  ACCCGGAGCGCTTCTCTCTGTGGCCAGGCTTCAACCAAGTTCAAGCTCC 1704
Qy      1879  ACATTGACCTCCGCAATGAGCTCCAGGTTTACAGAAATGCAAGAGCTGGGGTGG 1938
Db      1705  ACATTGACCTCCGCAATGAGCTCCAGGTTTACAGAAATGCAAGAGCTGGGGTGG 1764
Qy      1939  CAGGCAAGACTGCTTTTAAATGCTCCATTTCAAGAGATACACCGAGACTCGG 1998
Db      1765  CAGGCAAGACTGCTTTTAAATGCTCCATTTCAAGAGATACACCGAGACTCGG 1824
Qy      1999  AGGGAGACAGTAGACACCAAGGCCCACTTTGTCCCTTACGAATTCAGGTTACGCTC 2058
Db      1825  AGGGAGACAGTAGACACCAAGGCCCACTTTGTCCCTTACGAATTCAGGTTACGCTC 1884
Qy      2059  CACCTAGAACAGAGGCTGCTTACGATGCTGTTCTTCAAGCATTTATTAAGACCTAC 2118
Db      1885  CACCTAGAACAGAGGCTGCTTACGATGCTGTTCTTCAAGCATTTATTAAGACCTAC 1944
Qy      2119  TGGGTGCTGGGTCACTGTGCTTAGAAGAACAGAGGGTCCCAAGTCTTGGCTTGGC 2178

```

```

Db      1945  TGGGTGCTGGGTCACTGTGCTTAGAAGAACCAAGAGGTCCCAAGTCTTGGCTTGGC 2004
Qy      2179  GCGCCCTGCTGCCCCACCACTTTCTGCAACACAGCGGTGGGAGGCGGGAGAGCAGC 2238
Db      2005  GCGCCCTGCTGCCCCACCACTTTCTGCAACACAGCGGTGGGAGGCGGGAGAGCAGC 2064
Qy      2239  TGGGACCCAGAACTGAGCTGGGAGGATCCGACAGAAAAGCTCAGGGCGGGTCTTCC 2298
Db      2065  TGGGACCCAGAACTGAGCTGGGAGGATCCGACAGAAAAGCTCAGGGCGGGTCTTCC 2124
Qy      2299  TTGTCCTCCGGATTTGGCTTATGCTGGGTACCAACATGTAATCAGGATGTGGTTTGA 2358
Db      2125  TTGTCCTCCGGATTTGGCTTATGCTGGGTACCAACATGTAATCAGGATGTGGTTTGA 2184
Qy      2359  ACCCAATAACAAAGGCTTGTGCTCAGCTTTTACAGATATTTTGTATTTAACT 2418
Db      2185  ACCCAATAACAAAGGCTTGTGCTCAGCTTTTACAGATATTTTGTATTTAACT 2244
Qy      2419  CTCTAAGATATTTGATTTTATGAGGCTTAAAGAACTTATGATCTTATTTGGTCTT 2478
Db      2245  CTCTAAGATATTTGATTTTATGAGGCTTAAAGAACTTATGATCTTATTTGGTCTT 2304
Qy      2479  TCTGAGTTCAAGAGGCTTAACTTCTCAGCTCAACAGCAAGTCTGTGGTGGC 2538
Db      2305  TCTGAGTTCAAGAGGCTTAACTTCTCAGCTCAACAGCAAGTCTGTGGTGGC 2364
Qy      2539  AGAAGCAAGCTTAAAGGCTTAACTTCAAGTCAATCAAGATGCTCTCTTGAATGCT 2598
Db      2365  AGAAGCAAGCTTAAAGGCTTAACTTCAAGTCAATCAAGATGCTCTCTTGAATGCT 2424
Qy      2599  TGGCCCTGAGGCTTCAAGGCTTGGAGACTTGTGCTCAACCTCTCTCTAGATGAG 2658
Db      2425  TGGCCCTGAGGCTTCAAGGCTTGGAGACTTGTGCTCAACCTCTCTCTAGATGAG 2484
Qy      2659  TCTGTAGAGGCTCTGTAGATATTTGTGTAACACATGCCCATATATACAGTACACA 2718
Db      2485  TCTGTAGAGGCTCTGTAGATATTTGTGTAACACATGCCCATATATACAGTACACA 2544
Qy      2719  GATGTACACAGAGATGTAACATGCTCAGGCCCAAGCTCTGATACCTGACCTGACACC 2778
Db      2545  GATGTACACAGAGATGTAACATGCTCAGGCCCAAGCTCTGATACCTGACCTGACACC 2604
Qy      2779  CAGCTTGGCCCTGCTGCTGCTGCTGCTCAAGCAGAGCTCAACCTGCTCTGTCC 2838
Db      2605  CAGCTTGGCCCTGCTGCTGCTGCTGCTCAAGCAGAGCTCAACCTGCTCTGTCC 2664
Qy      2839  CCTTCCCAACCACTGCTGAGCTTCTGAGCAGACAGTACCTTGGCTGACCGGTGT 2898
Db      2665  CCTTCCCAACCACTGCTGAGCTTCTGAGCAGACAGTACCTTGGCTGACCGGTGT 2724
Qy      2899  GTGGCCGCTCTCAACCGAGGACAGCCCGGACCAATGGAATCTCGGTGACATCAAT 2958
Db      2725  GTGGCCGCTCTCAACCGAGGACAGCCCGGACCAATGGAATCTCGGTGACATCAAT 2784
Qy      2959  AAAAGTGGGTTTGTATCAAAAAA 3006
Db      2785  AAAAGTGGGTTTGTATCAAAAAA 2832

```

RESULT 6  
 ABX72001  
 ID ABX72001 standard; DNA; 2832 BP.  
 XX  
 AC ABX72001;  
 XX  
 DT 12-MAR-2003 (first entry)  
 XX  
 DE DNA encoding human tumour endothelial marker TEM 2.  
 XX  
 KW Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
 KW Tumour endothelial marker; normal endothelial marker; PEM;  
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;

KW neovascularization; immune response; cytoskeletal; antidiabetic; gene;  
 KW ophthalmological; antineoplastic; antirheumatic; antidiabetic; antiproliferative; ds.  
 OS Homo sapiens.  
 PN WO200283874-A2.  
 PD 24-OCT-2002.  
 PP 10-APR-2002; 2002WO-US008253.  
 PR 11-APR-2001; 2001US-0282850P.  
 PR 06-FEB-2002; 2002US-0354262P.  
 XX (UJJO) UNIV JOHNS HOPKINS.  
 XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;  
 DR MPI; 2003-093016/08.  
 DR P-PSDB; ABUS4429.  
 XX New purified human transmembrane protein, designated as tumor endothelial  
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,  
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.  
 XX  
 XX Disclosure; Page 113-114; 374pp; English.  
 XX The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumor ECs. Tumor endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumors as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neovascularization or tumor angiogenesis, for  
 CC inducing an immune response to tumor endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumors. The present  
 CC sequence represents a human TEM or NEM gene of the invention  
 XX  
 SQ Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 U; 0 Other;  
 Query Match 91.1%; Score 2806.4; DB 7; Length 2832;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 385 CTGAGGTCAAGTCTGCTGTAAGAACCAAGAGAGGCGGAGTCCATGCTC 444  
 QY ATCTGTGGCAACAGAACAGACCGGAGGCTGTCGCGCCAGTCCCAACACGAGGCC 678  
 DB 445 ATCTGTGGCAACAGAACAGACCGGAGGCTGTCGCGCCAGTCCCAACACGAGGCC 504  
 QY GAGCTGCTGATGTGGGCGAGAGAACTCCGCTACTTCAAGTGTCCGACCAAGAAAG 738  
 DB 505 GAGCTGCTGATGTGGGCGAGAGAACTCCGCTACTTCAAGTGTCCGACCAAGAAAG 564  
 QY 739 ACCAAGTGAAGAGATGTTACGTGCTCTTACAGATGACCAAGTCCCAACAGATG 798  
 DB 565 ACCAAGTGAAGAGATGTTACGTGCTCTTACAGATGACCAAGTCCCAACAGATG 624  
 QY 799 AGCCCCGCTGATCGGAAGATCTCCGTGAGATGAGTACGCTTCCACCCAGAGGCC 858  
 DB 625 AGCCCCGCTGATCGGAAGATCTCCGTGAGATGAGTACGCTTCCACCCAGAGGCC 684  
 QY TTCTGATGCGCGCGGTCAAGAGATGAGCGCTATGAGATGTCGCGCTTCCGCGCC 918  
 DB 685 TTCTGATGCGCGCGGTCAAGAGATGAGCGCTATGAGATGTCGCGCTTCCGCGCC 744  
 QY 919 CGCCCAAGCTCAACAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 978  
 DB 745 CGCCCAAGCTCAACAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 804  
 QY 979 GCGCGTGAAGAGAGACAGTGAAGTCAAGTGAAGAGAGATGCTGGGGCGGGCTTGGC 1038  
 DB 805 GCGCGTGAAGAGAGACAGTGAAGTCAAGTGAAGAGAGATGCTGGGGCGGGCTTGGC 864  
 QY 1039 CAGTGCCTTCAAGGAGAGTGGGCGGAGATGCGCATGTCGATGTCGCGCATGTCGCGCAT 1098  
 DB 865 CAGTGCCTTCAAGGAGAGTGGGCGGAGATGCGCATGTCGATGTCGCGCATGTCGCGCAT 924  
 QY 1099 GCGAGCACTTCTTCTTCAAGACCTTGAAGACCAAGTGAAGAGAGAGAGAGAGAGAG 1158  
 DB 925 GCGAGCACTTCTTCTTCAAGACCTTGAAGACCAAGTGAAGAGAGAGAGAGAGAGAG 984  
 QY 1159 GCGACATTCCTGCTGCTTCTTCAAGACCTTGAAGACCAAGTGAAGAGAGAGAGAG 1218  
 DB 985 GCGACATTCCTGCTGCTTCTTCAAGACCTTGAAGACCAAGTGAAGAGAGAGAGAG 1044  
 QY 1219 GTTTCATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1278  
 DB 1045 GTTTCATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104  
 QY 1279 AGTTACACCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1338  
 DB 1105 AGTTACACCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1164  
 QY 1339 TGCCTCATATGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1398  
 DB 1165 TGCCTCATATGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1224  
 QY 1399 GAGAAAGGTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458  
 DB 1225 GAGAAAGGTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1284  
 QY 1459 ACTTCTCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1518  
 DB 1285 ACTTCTCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344  
 QY 1519 ACCCCAGACACACACAAAGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1578  
 DB 1345 ACCCCAGACACACACAAAGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1404  
 QY 1579 GTGCTGAAGGGGGGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1638  
 DB 1405 GTGCTGAAGGGGGGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1464  
 QY 1639 CCTCTGAAGGGGGGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1698  
 DB 1465 CCTCTGAAGGGGGGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1524





```

FT variation replace(10292,G)
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(10792,G)
FT /*tag= n
FT /standard_name= "Single nucleotide polymorphism (SNP)"
PN MO200262849-A2.
XX 15-AUG-2002.
XX
XX 28-JAN-2002; 2002WO-US002301.
XX
XX 08-FEB-2001; 2001US-00778963.
XX
XX (PEKE ) PE CORP NY.
XX
XX Neelam B, Ketchum KA, Di Francesco V, Beasley EM;
XX
XX MPI; 2002-706901/76.
XX
XX P-PSDB; ABG70951.
XX
XX New human Ras-like polypeptides, useful for treating diseases associated
XX with an increase in apoptosis or cell proliferation e.g., cancer.
XX
XX Claim 4; Fig 3; 73pp; English.
XX
XX The present invention relates to a new polypeptide comprising a fully
XX defined 266-amino acid sequence or its fragment, orthologue or allelic
XX variant. The allelic variant or orthologue is encoded by a nucleic acid
XX molecule that hybridises under stringent conditions to the opposite
XX strand of a nucleic acid molecule with a fully defined sequence of 3082
XX or 11221 bp. The pharmaceutical composition comprising the agent that
XX binds to the polypeptide is useful for treating a disease or condition
XX mediated by a human Ras-like protein e.g., cancer. Likewise, the
XX polypeptide is useful for identifying modulators of Ras-like protein
XX activity or expression by screening a compound for the ability to
XX stimulate or inhibit interaction between the Ras-like protein and an
XX interacting molecule or Ras-like activity. The polypeptide is useful for
XX diagnosing a disease, or predisposition to a disease, or treating
XX diseases associated with an increase in apoptosis or cell proliferation
XX e.g., AIDS (acquired immunodeficiency syndrome), Alzheimer's disease,
XX Parkinson's disease, aplastic anaemia, allergies, atherosclerosis,
XX osteoporosis or bacterial, viral, fungal, parasitic, protozoal or
XX helminthic infections, and inflammation. The present nucleic acid
XX sequence encodes the human Ras-like protein of the invention
XX
XX Sequence 11221 BP; 2512 A; 3108 C; 3028 G; 2573 T; 0 U; 0 Other;
XX
XX Query Match 81.1%; Score 2500.4; DB 6; Length 11221;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 2504; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 769 TTCAGCATGCGCAAGCTGCCACAGAGATGAGCCCGCCCTGCAATCGCAAGATCTCCG 828
DB 7981 TTCAAGATGCGCAAGCTGCCACAGAGATGAGCCCGCCCTGCAATCGCAAGATCTCCG 8040
QY 829 CAGTACGGTGAAGCGCTTCCACACCCAGAGCGCTTTCGATCGCCCGCTCAAGAGATGAG 888
DB 8041 CAGTACGGTGAAGCGCTTCCACACCCAGAGCGCTTTCGATCGCCCGCTCAAGAGATGAG 8100
QY 889 GCCTATGCGATGATGCTGCGCCCTTGGCCCGCCCGCCAGAGCTTCAAGTACCTCAAGTAC 948
DB 8101 GCCTATGCGATGATGCTGCGCCCTTGGCCCGCCCGCCAGAGCTTCAAGTACCTCAAGTAC 8160
QY 949 ATCAAGGCGCAAGTCTTCCGAGGAAGCCAGGCGCTGAGAGGAGCAAGTGAACCTTCAG 1008
DB 8161 ATCAAGGCGCAAGTCTTCCGAGGAAGCCAGGCGCTGAGAGGAGCAAGTGAACCTTCAG 8220
QY 1009 TGAGCGAGGATGCTGCGCGCGCGCTGGCCAGTGCCTTCAAGGAGGTGCGCCCGAGATGC 1068
DB 8221 TGAGCGAGGATGCTGCGCGCGCGCTGGCCAGTGCCTTCAAGGAGGTGCGCCCGAGATGC 8280
QY 1069 CCATGTGCGCATCTCCCGACCGAGCGCCCGCCAGAGCTTGTTCACAGACTTATGCGCA 1128
DB 8281 CCATGTGCGCATCTCCCGACCGAGCGCCCGCCAGAGCTTGTTCACAGACTTATGCGCA 8340
QY 1129 CCAGACTGAGAGCGCCCGCGCGCGCTGGCCAGTGCCTTCAAGGAGGTGCGCCCGAGATGC 1188
DB 8341 CCAGACTGAGAGCGCCCGCGCGCGCTGGCCAGTGCCTTCAAGGAGGTGCGCCCGAGATGC 8400
QY 1189 CTGAGTCCGCTTGTTCACAGCTCTTGTGATGATCTCTCTGTGAGAGACATCT 1248
DB 8401 CTGAGTCCGCTTGTTCACAGCTCTTGTGATGATCTCTCTGTGAGAGACATCT 8460
QY 1249 CTGAGCTCTCAAGATGAGGAGAGACTCAAGTACACTTCTCTCTGTGAGAGAG 1308
DB 8461 CTGAGCTCTCAAGATGAGGAGAGACTCAAGTACACTTCTCTCTGTGAGAGAG 8520
QY 1309 AATGTTGATCCAGAGGAGGAGAGATGAGTGTGCTGATGATGAGAGCTCTGAGAGCAAGCC 1368
DB 8521 AATGTTGATCCAGAGGAGGAGAGATGAGTGTGCTGATGATGAGAGCTCTGAGAGCAAGCC 8580
QY 1369 TCAGATGATAAAGAGACAGAGAGGCGCAGATGAGAAAGTCTCTCTCTGTGAGATACA 1428
DB 8581 TCAGATGATAAAGAGACAGAGAGGCGCAGATGAGAAAGTCTCTCTCTGTGAGATACA 8640
QY 1429 CCCAGCTGTTGTTGGTGTGAGCGCTGAGAGAACTTCTCTCCAGCGCTTCACTTAAAGC 1488
DB 8641 CCCAGCTGTTGTTGGTGTGAGCGCTGAGAGAACTTCTCTCCAGCGCTTCACTTAAAGC 8700
QY 1489 TCTGATTCAGCTGCTGTCAGACCCCTCCAGCCCGCCAGACACACAAGTTGAGCCCGCA 1548
DB 8701 TCTGATTCAGCTGCTGTCAGACCCCTCCAGCCCGCCAGACACACAAGTTGAGCCCGCA 8760
QY 1549 GCTGCGCTGACATTTGAGCCAGTGAATCTGTGTCTGAAGGAGGAGGAGGAGCAACCTTCT 1608
DB 8761 GCTGCGCTGACATTTGAGCCAGTGAATCTGTGTCTGAAGGAGGAGGAGGAGCAACCTTCT 8820
QY 1609 AGACCAAGCCCAACCACTTAAACACAGCCCACTCTGACCGCGTCTCTCAAGCTCTCTC 1668
DB 8821 AGACCAAGCCCAACCACTTAAACACAGCCCACTCTGACCGCGTCTCTCTCAAGCTCTCTC 8880
QY 1669 CTAAGTCCCTCCGCGCCAGCACTTGTGCTTGTGAGTGAAGCTGTTTGTGTGATGT 1728
DB 8881 CTAAGTCCCTCCGCGCCAGCACTTGTGCTTGTGAGTGAAGCTGTTTGTGTGATGT 8940
QY 1729 ATAGTAGTAATGAAATGAAATGTAATCTGTAAAGCTAGTGAATCTCTCTTGTGAGCAG 1788
DB 8941 ATAGTAGTAATGAAATGAAATGTAATCTGTAAAGCTAGTGAATCTCTCTTGTGAGCAG 9000
QY 1789 GCCCTACCCAGATTCAATTCAGAGCGCTTCCACCCGCGAGAGCGCTTCTCTCTCTCCAA 1848
DB 9001 GCCCTACCCAGATTCAATTCAGAGCGCTTCCACCCGCGAGAGCGCTTCTCTCTCTCCAA 9060
QY 1849 ACAGGATTCGATGCTGTTTGTGAGCTGAGATGAGCTCTCCGCAATGAGCTCCACGGT 1908

```

Dp	9061	ACAGGGTTCCGTGGCCCTGTTTGACGTGACATTGA	CTCCGCGCATTTAGAGCTCCAGGT	9120
Qy	1909	TTACAGACAATTGCAACAAGCGTGGGGTGGAGGCG	CAGACATGCTTTTATTAATGCTCC	1968
Dp	9121	TTACAGACAATTGCAACAAGCGTGGGGTGGAGGCG	CAGACATGCTTTTATTAATGCTCC	9180
Qy	1969	CATTTCAACAGAGGATACCAACCGAGCTCGAGGGG	AGACACATGAGCACAGGCCCACT	2028
Dp	9181	CATTTCAACAGAGGATACCAACCGAGCTCGAGGGG	AGACACATGAGCACAGGCCCACT	9240
Qy	2029	TTGTCCCTAGCAAAATTGACGGGTACAGTCCAC	CTAGAACAGAGCTGCGCTCTACTGTGC	2088
Dp	9241	TTGTCCCTAGCAAAATTGACGGGTACAGTCCAC	CTAGAACAGAGCTGCGCTCTACTGTGC	9300
Qy	2089	TGCTTCTCMAAGCATTTATTAAGCACTTAC	TGGGTGCTGGGTCACTGTGTCTTAGGAAA	2148
Dp	9301	TGCTTCTCMAAGCATTTATTAAGCACTTAC	TGGGTGCTGGGTCACTGTGTCTTAGGAAA	9360
Qy	2149	CCAAAGGGGTCCCGAGCTCGGAGCTCGGCGCCG	CTGCGCCACACCTTCGTGACA	2208
Dp	9361	CCAAAGGGGTCCCGAGCTCGGAGCTCGGCGCCG	CTGCGCCACACCTTCGTGACA	9420
Qy	2209	CACAGCGGTGGGAGGCGGGAGGAGCAGCTGGAG	CCAGAACCTAGACCTGGAGGATC	2268
Dp	9421	CACAGCGGTGGGAGGCGGGAGGAGCAGCTGGAG	CCAGAACCTAGACCTGGAGGATC	9480
Qy	2269	CGACGAAAACTCAGAGGCGGGGTCTTCTCTGT	GCCCGGATTTGCGGTATGCTGGGTAC	2328
Dp	9481	CGACGAAAACTCAGAGGCGGGGTCTTCTCTGT	GCCCGGATTTGCGGTATGCTGGGTAC	9540
Qy	2329	CACCATGTACTCAGGCACTGTGGGTTTGAACCC	ATTAACCAAAAGGCCCTTGTCATCAGC	2388
Dp	9541	CACCATGTACTCAGGCACTGTGGGTTTGAACCC	ATTAACCAAAAGGCCCTTGTCATCAGC	9600
Qy	2389	TCTTTAACAGATATATTTTGTATTTTAACTCT	CTCTAATAATTTGAAGTTTATGGGCCCTTA	2448
Dp	9601	TCTTTAACAGATATATTTTGTATTTTAACTCT	CTCTAATAATTTGAAGTTTATGGGCCCTTA	9660
Qy	2449	AGGAACCTTAGTGATCTTCTATTGGGACTTT	CTAGAGTTGAGAGGGGTAAATGTA	2508
Dp	9661	AGGAACCTTAGTGATCTTCTATTGGGACTTT	CTAGAGTTGAGAGGGGTAAATGTA	9720
Qy	2509	TCCAAGTTCACACAGCAAGCTGTGGGTGGGAG	AGAAAGCAAGCTAGCGCTAGTCAAGTAC	2568
Dp	9721	TCCAAGTTCACACAGCAAGCTGTGGGTGGGAG	AGAAAGCAAGCTAGCGCTAGTCAAGTAC	9780
Qy	2569	ATTACCAAGATGTGCTCCCTCTCTTGATGCTTT	GGCGCCCTTCAGAGGCTTTGGAGAC	2628
Dp	9781	ATTACCAAGATGTGCTCCCTCTCTTGATGCTTT	GGCGCCCTTCAGAGGCTTTGGAGAC	9840
Qy	2629	ATCTTGTCTCAACCTCTCCCTAGATCAGTCTGT	AGAGGCTCCCTGTAGATATTGTGTAC	2688
Dp	9841	ATCTTGTCTCAACCTCTCCCTAGATCAGTCTGT	AGAGGCTCCCTGTAGATATTGTGTAC	9900
Qy	2689	ACCATGCCAATGTATATTAACAAGTACACACAG	ATGTATACACATGTCTTCAG	2748
Dp	9901	ACCATGCCAATGTATATTAACAAGTACACACAG	ATGTATACACATGTCTTCAG	9960
Qy	2749	CCCCAGCTCTGCATACCTGCAACCTGCAACCC	CCAGGCTTGGCGCCCTGCTGTGTGTAC	2808
Dp	9961	CCCCAGCTCTGCATACCTGCAACCTGCAACCC	CCAGGCTTGGCGCCCTGCTGTGTGTAC	10020
Qy	2809	AAAGCAGAGCTCCAAACCTGTGCTGTCTCCCT	TCCCAACCACTGTAGAGCTTCTGA	2868
Dp	10021	AAAGCAGAGCTCCAAACCTGTGCTGTCTCCCT	TCCCAACCACTGTAGAGCTTCTGA	10080
Qy	2869	GCAGACAGAGTACTTGGCTGTACACCGGTGTGT	GGCCGCTACACCAAGGACAGGCCGG	2928
Dp	10081	GCAGACAGAGTACTTGGCTGTACACCGGTGTGT	GGCCGCTACACCAAGGACAGGCCGG	10140
Qy	2929	CCACCATGATCTTCGTGTACATATCAATTAAG	TGTGGTTGTACA	2978

D	b
<hr/>	
10141 CCACCATGAGTCCTCCGTGTACACTCATCAATAAAGAATGGTTTGTAACMA 10190	
<hr/>	
AAZ36913	RESULT 8
ID AAZ36913	standard; DNA; 837 BP.
XX AC AAZ36913;	
XX DT 13-MAR-2000	(first entry)
DE XX	DNA encoding a homologue of activator of G protein signalling AGS1.
KW KM	Activator of G protein signalling; AGS; ras-related G protein; GRP hydrolysis; G protein activity; pheromone response pathway; KW KW cellular-coupled signal transduction; G-gamma selectivity; cellular signal transduction; AGS1 homologue; ss. XX OS Homo sapiens. XX FH Key Location/Qualifiers FH CDS 1..837 /tag= a FT FT /product= "AGS1 homologue"
PX MO9358670-AI.	
PN 18-NOV-1999.	
PF PD 07-MAY-1999;	99WO-USO10151.
PR PR 08-MAY-1998:	98US-0084842P.
PP PR 07-OCT-1998;	98US-0103355F.
PA CADU-) CADUS PHARM CORP.	
PI Cismowski M,	Duzic E;
DR WPI: 2000-072337/06.	
DR P-PsDB; AA53923.	
PT A new activator of G protein signalling used to treat disorders	
PS characterized by an aberrant AGS protein activity.	
XX Example 15; Page 143-144; 162pp; English.	
CC The present sequence encodes a homologue of a human AGS1 (activator of G	
CC protein signalling (AGS)) protein. The AGS cDNA sequence was isolated	
CC from a human liver cDNA library. The AGS protein exhibits homology to ras-	
CC related G proteins, and contains alterations in conserved amino acids	
CC consistent with a deficiency in GRP hydrolysis activity. AGS stimulates G	
CC protein activity, G protein-coupled signal transduction and the pheromone	
CC response pathway in a receptor-independent manner. The AGS protein also	
CC shows G-gamma selectivity, as measured by growth assays in yeast	
CC expressing various mammalian G-gamma constructs, and tissue-specific	
CC expression, as measured by Northern blot analysis. The AGS protein can be	
CC used to screen for compounds that modulate cellular signal transduction.	
CC The protein is used to treat disorders characterized by an aberrant AGS	
CC protein activity or AGS nucleic acid expression	
SQ Sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 U; 0 Other;	
Query Match	26.3%; Score 811.4; DB 3; Length 837;
Best Local Similarity	99.9%; Pred. No. 2.2e-111;
Matches 812; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
OY 199 CAGCCCCCAGGCATGATGAAGACTTTCGCCAGCGGAGAACGTGACGCTCAGTGTGCCC GCC 258	
Db 25 CAGCCCCCAGGCATGATGAAGACTTTCGCCAGCGGAGAACGTGACGCTCAGTGTGCCC GCC 84	
OY 259 AAAAATTATATCCGATGATGTGTGTGCTTCCGTGCGTAGGAGAGAGCTCATTCGTG 318	
Db 85 AAAAATTATATCCGATGATGTGTGTGCTTCCGTGCGTAGGAGAGAGCTCATTCGTG 144	

QY	31.9	TCCTGGCTTCCTCAAGTGGCGGCTTTGAGGACCAAGTACACACCCACCAATCCGAGGACTTCCAC	37.8
Db	145	TTCTCGCTTCTCTTAATGAGCGGCTTTGAGGACCAAGTACACACCCACCAATCCGAGGACTTCCAC	204
QY	37.9	CGTAAGGATATACAACTCCGCGGCGCATGTACACAGCTCGACATCTCTGTATACCTCTGGC	43.8
Db	205	CGTAAGGATATACAACTCCGCGGCGCATGTACACAGCTCGACATCTCTGTATACCTCTGGC	264
QY	43.9	AACCAACCCCTTCCCGGCGCATGCGCAGGCGTGCATCTCAAGGGATAGTCTTCAATCTCG	4.98
Db	265	AACCAACCCCTTCCCGGCGCATGCGCAGGCGTGCATCTCAAGGGATAGTCTTCAATCTCG	3.24
QY	49.9	GTGTTCAAGCTGTGATTAACCGGAGAGTCTTGCATGAGGTCAAGCGCCTTCAGAGCAGATC	55.8
Db	325	GTGTTCAAGCTGTGATTAACCGGAGAGTCTTGCATGAGGTCAAGCGCCTTCAGAGCAGATC	38.4
QY	55.9	CTGAGAGGTAAAGTCTCTGCTTGAAGAAACAAGACCAAGAGGCGGCGGAGCTGCCATAGTGC	61.8
Db	385	CTGAGAGGTAAAGTCTCTGCTTGAAGAAACAAGACCAAGAGGCGGCGGAGCTGCCATAGTGC	44.44
QY	61.9	ATCTGTGGCAACAAGACGACCAACCGCGAGCTGTGCGCGCAGGTGCCACACCGAGGCC	67.87
Db	445	ATCTGTGGCAACAAGACGACCAACCGCGAGCTGTGCGCGCAGGTGCCACACCGAGGCC	50.4
QY	67.9	GAGCTGCTGCTGTCCGGGCGACGAGAACTCCGCTTACTTGAGGTTGTGCGCCAAAGAAAC	73.8
Db	505	GAGCTGCTGCTGTCCGGGCGACGAGAACTCCGCTTACTTGAGGTTGTGCGCCAAAGAAAC	56.4
QY	73.9	ACCAACGTGGAGAGAGATGTTTCAACGTGCTTTCAGATGGCCAAAGCTGCCACAGAGATG	79.8
Db	555	ACCAACGTGGAGAGAGATGTTTCAACGTGCTTTCAGATGGCCAAAGCTGCCACAGAGATG	62.4
QY	79.9	AGCCCGCGCTTCATTCGACAAGATCTTCGTGACATGACGTGACGCTTCCACCCACGAGCC	85.8
Db	625	AGCCCGCGCTTCATTCGACAAGATCTTCGTGACATGACGTGACGCTTCCACCCACGAGCC	68.4
QY	85.9	TTCTGCATGCGCGCGCTCAAGAGAGATGGACGCTTAATGGCATGGTTCGCGCCTTCGCGCG	91.8
Db	685	TTCTGCATGCGCGCGCTCAAGAGAGATGGACGCTTAATGGCATGGTTCGCGCCTTCGCGCG	74.44
QY	91.9	CGCCCGCAGCGTCAACAGTGAAGCTTCAATCAATCAAGGCGCAAGGTCTCTTCCGAGAGGCGCAG	97.87
Db	745	CGCCCGCAGCGTCAACAGTGAAGCTTCAATCAATCAAGGCGCAAGGTCTCTTCCGAGAGGCGCAG	80.4
QY	97.9	GCCCGTGAGAGGAGCAAGTGCACATCCCACTGTA	101.1
Db	805	GCCCGTGAGAGGAGCAAGTGCACATCCCACTGTA	83.7

Query Match	Best Local Similarity	23.6%	Score 728.8	DB 5	Length 951
Matches 760	Conservative	0	Mismatches 52	Indels 0	Gaps 0
269	ACCGCATGATGATGCTGGGATCCTTCGCGATGGGCAAGAGCTCCATCGTGTCTCGCTTC	328			
140	AACGATGATGATGCTGGGATCCTTCGCGATGGGCAAGAGCTCCATCGTGTCTCGCTTC	199			
329	TCAATGCGCGCTTTGAGGACCAATGACACCCACCATGAGGACTTCCACCGTAAGAT	388			
200	TCAATGCGCGCTTTGAGGACCAATGACACCCACCATGAGGACTTCCACCGTAAGAT	259			
389	ACAATATCGGGGCGACATGTAACAAGCTGCAATCTGGAATACCTTGGCAACCAACCCCT	448			
260	ACAATATCGGGGCGACATGTAACAAGCTGCAATCTGGAATACCTTGGCAACCAACCCCT	319			
449	TCCCGCGCATGCGAGGCTGTCCATCCTCAAGGAGGATGTCTTCACTCTGATTCAGCC	508			
320	TCCCGCGCATGCGAGGCTGTCCATCCTCAAGGAGGATGTCTTCACTCTGATTCAGCC	379			
509	TGATTAACCGGAGTCTTTCATGATGATCAAGCGCTTCAAGAGCAATCTTGAAGGTCA	568			
380	GCGGAGGCGCAAGGCGATGGGTGGGAGATGTGCTGGGACCTTAGCAGATCTCGAGGTCA	439			
569	AGTCTGCTTAAGAAACAAGCAAGAGGCGGCGGAGCTGCCATGATCATCTGTGGCA	628			
440	AGTCTGCTTAAGAAACAAGCAAGAGGCGGCGGAGCTGCCATGATCATCTGTGGCA	499			
629	ACAAGAACGACCAAGGCGAGCTGTGTCGCCAGAGTGGCCACCAAGGCGGAGCTGTG	688			
500	ACAAGAACGACCAAGGCGAGCTGTGTCGCCAGAGTGGCCACCAAGGCGGAGCTGTG	559			
689	TGTCGGGCGACGAGAACTTCGCGCTTCTTCAAGGTGTGGCCAGAGAAACACCAAGTGG	748			

Db 560 TGTGGGGGACGAGACTGGCTACTTCCAGGTGTGGCGAAGAAACACCAAGCTGG 619  
 QY 749 ACGAGATGTTCTACGCTCTTTCAGCATGGCCAAAGTGCACAGAGATGAGCCCGCC 808  
 Db 620 ACGAGATGTTCTACGCTCTTTCAGCATGGCCAAAGTGCACAGAGATGAGCCCGCC 679  
 QY 809 TGGATGCAAGATCTCCGTCAGTGAAGGTGAGCCCTTCCACCCCAAGCCCTTTCGATG 868  
 Db 680 TGCATGCAAGATCTCCGTCAGTGAAGGTGAGCCCTTCCACCCCAAGCCCTTTCGATG 739  
 QY 869 GCCGCTCAAGAGATGAGAGCTTATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 928  
 Db 740 GCCGCTCAAGAGATGAGAGCTTATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 799  
 QY 929 TCACAGTGAAGTCAAGTACATCAAGGCGCAAGGTCCTTGGGAGAGCCAGCCCTGAGA 988  
 Db 800 TCACAGTGAAGTCAAGTACATCAAGGCGCAAGGTCCTTGGGAGAGCCAGCCCTGAGA 859  
 QY 989 GGGACAGTGCACCATCCAGTGAAGGAGATGCTGGGGCGGGGCTTGGCCAGTGCCTTC 1048  
 Db 860 GGGACAGTGCACCATCCAGTGAAGGAGATGCTGGGGCGGGGCTTGGCCAGTGCCTTC 919  
 QY 1049 AGGAGGTGGCCCGCCAGATGCCCACTGTGCGCA 1080  
 Db 920 AGGAGGTGGCCCGCCAGATGCCCACTGTGCGCA 951

## RESULT 10

ABL92134  
 ID ABL92134 standard; cDNA; 3020 BP.

AC ABL92134;

XX 30-MAY-2002 (first entry)

DE Mouse Tumour Endothelial Marker polynucleotide SEQ ID NO 292.

XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytosstatic;

KW normal endothelial marker; pan-endothelial marker; immunostimulant;

KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;

KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;

XX psoriasis; gene; sb.

OS Mus musculus.

PN WO200210217-A2.

PD 07-FEB-2002.

PF 01-AUG-2001; 2001WO-US024031.

PR 02-AUG-2000; 2000US-0222599P.

PR 11-AUG-2000; 2000US-0224360P.

PR 11-APR-2001; 2001US-0282850P.

PA (UYUO) UNIV JOHNS HOPKINS.

PI St Croix B, Kinzler KW, Vogelstein B;

DR WPI; 2002-291856/33.

DR P-PSDB; ABB90781.

XX An isolated molecule comprising an antibody variable region which

PT specifically binds to an extracellular domain of a tumor endothelial

XX marker (TEM) protein, useful for inhibiting tumor growth.

PS Disclosure; Page 294-295; 331pp; English.

CC The invention relates to an isolated molecule comprising an antibody

CC variable region which specifically binds to an extracellular domain of a

CC tumor endothelial marker (TEM) protein selected from ABB90732, ABB90740,

CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM

CC proteins have cytosstatic, immunostimulant and antiangiogenic activity.

CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects

CC bearing a vascularised tumour, polycystic kidney disease, diabetic

CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM

CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90769)

CC are disclosed, as are marker oligonucleotide sequences: tumour

CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal

CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers

CC (PEM) ABL91903-ABL91995

XX

SQ Sequence 3020 BP; 690 A; 848 C; 780 G; 702 T; 0 U; 0 Other;

Query Match 22.5%; Score 693.2; DB 6; Length 3020;

Best Local Similarity 87.0%; Pred. No. 7.4e-94;

Matches 774; Conservative 0; Mismatches 113; Indels 3; Gaps 1;

QY 200 AGCCCGAGCCATGATGAAGACTTGTTCAGCGGGAACTGACGCTCATGTGGCCCGCA 259

Db 365 AGCCAGAGCCATGATGAAGACTTGTTCAGCGGGAACTGACGCTCATGTGGCTGCTA 424

QY 260 AAAATCTATACCGCATGT 319

Db 425 AGAATCTCTACCGCATGT 484

QY 320 CTGCTTCTCTAATGTGGCGCTTTGAGAGACAGTACACCCACATCGAGACTTCCACC 379

Db 485 CCGCTTCTCTAATGTGGCGCTTTGAGAGACAGTACACCCACATCGAGACTTCCACC 544

QY 380 GTAAGGTATACATCCCGCGCGCATGTATCAAGCTTCATCTGTATCTCTGCA 439

Db 545 GCAAGGTATACATCCCGCGCGCATGTATCAAGCTTCATCTGTATCTCTGCA 604

QY 440 ACCACCCCTTCCCGCGCATGT 499

Db 605 ACCACCCCTTCCCGCGCATGT 664

QY 500 TGTTCAGGCTGATTAACCGGAGTCTCTCATGTAGGTCAAGCGCTTCAGAAAGAGATCC 559

Db 665 TGTTCAGGCTGATTAACCGGAGTCTCTCATGTAGGTCAAGCGCTTCAGAAAGAGATCC 724

QY 560 TGAAGTCAAGTCTGTGCTGAAGAAACAAGCAAGAGCGCGAGACTGCTCCATGTGCA 619

Db 725 TGAAGTCAAGTCTGTGCTGAAGAAACAAGCAAGAGCGCGAGACTGCTCCATGTGCA 784

QY 620 TCTGTGGAACAACAAGCAAGCAAGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679

Db 785 TCTGTGGAACAACAAGCAAGCAAGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844

QY 680 AGCTGT 739

Db 845 AGCTGT 904

QY 740 CCAACGTGAGACAGATGTTCTTCACTGCTTTCAGCATGCGCAAGCTGCCACAGAGTGA 799

Db 905 CTAATGTGAACGAATGTTCTATGTGCTGTTCAGCATGCGCAAGCTGCCCATGAGATGA 964

QY 800 GCCCGCCCTGCAATCGCAAGATCTCGGTGAGTACGCTTCCACCCCAAGGCTT 859

Db 965 GCCCTGCACTGCACTAAGATCTCGGTGAGTACGCTTCCACCCCAAGGCTT 1024

QY 860 TCTGATGCGCGCGGTCAAGAGATGAGCGCTTATGAGCATGTGCTGCTGCTGCTGCTGCTG 919

Db 1025 TCTGATGCGCGCGGTCAAGAGATGAGCGCTTATGAGCATGTGCTGCTGCTGCTGCTGCTG 1084

QY 920 GCCCGCAAGTCAACAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 979

Db 1085 GCCCGCAAGTCAACAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1144

QY 980 CCCGTGAGAGGAGCAAGTGCACCATCAGTGA---GGAGAGATGTGTGGGGGGGGCTTG 1036

Db 1145 CCCGAGAGAGGAGCAAGTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1204

QY 1037 GCCAGTCTTTCAGGAGGTGGCCCAAGATGCCCACTGTGCGCATCTCC 1086

Db 1205 GCGAGTGCCTTAAGGAGGTGCTGAGGATTCCAGTCGCGACATCCAC 1254

RESULT 11

ABX72059

ID ABX72059 standard; DNA; 3020 BP.

XX

XX

AC ABX72059;

XX

DT 12-MAR-2003 (first entry)

XX

DE DNA encoding mouse tumour endothelial marker mTEM 2.

XX

XX Mouse; endothelial cell; EC; tumour endothelial cell; TEM; NEM;

KW Tumour endothelial marker; normal endothelial marker; PEM;

KW par-endothelial marker; polycystic kidney disease; psoriasis;

KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;

KW neovascularization; immune response; cytoskeletal; antidiabetic; gene;

KW ophthalmological; antineoplastic; antirheumatic; antipsoriatic; ds.

XX

OS Mus sp.

XX

PN WO200283874-A2.

XX

PD 24-OCT-2002.

XX

PF 10-APR-2002; 2002WO-US008253.

XX

PR 11-APR-2001; 2001US-0282850P.

PR 06-FEB-2002; 2002US-0354262P.

XX

PA (UYJO ) UNIV JOHNS HOPKINS.

XX

PI Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;

XX

DR WPI; 2003-093016/08.

DR P-PSDB; ABUS4488.

XX

PT New purified human transmembrane protein, designated as tumor endothelial

PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,

PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or

PT psoriasis.

XX

PS Disclosure; Page 329-331; 374pp; English.

XX

CC The present invention relates to a novel method for the isolation of

CC endothelial cells (ECs), and the identification of genes expressed in

CC normal and tumour ECs. Tumour endothelial marker (TEM), normal

CC endothelial marker (NEM), and par-endothelial marker (PEM) genes are

CC identified in human ECs. The human EC marker proteins and the

CC polynucleotide sequences encoding them are useful for detecting,

CC diagnosing or treating tumors as well as polycystic kidney disease,

CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also

CC useful for inhibiting neovascularization or tumour angiogenesis, for

CC inducing an immune response to tumour endothelial cells in a patient, or

CC for identifying candidate drugs for treating tumours. The present

CC sequence represents a mouse TEM gene

XX

SQ Sequence 3020 BP; 690 A; 848 C; 780 G; 702 T; 0 U; 0 Other;

XX

Query Match 22.5%; Score 693.2; DB 7; Length 3020;

Best Local Similarity 87.0%; Pred. No. 7.4e-94;

Matches 774; Conservative 0; Mismatches 113; Indels 3; Gaps 1;

QY 200 AGCCCGAGCCATGATGAAGACTTTGTCACGCGGAGATGACGCTCATGTGCGCCGCA 259

Db 365 AGCCGAGGCGCATGATGAAGACTTTGTCACGCGGAGATGACGCTCATGTGCGCTA 424

QY 260 AAAACTCATACCGCATGTGTGCTGCGCTCTCGGAGTGGCAAGCTCCATCGTGT 319

Db 425 AGAAGTCTTACCGCATGTGTGCTGCGCTCTCGGAGTGGCAAGCTCCATCGTGT 484

QY 320 CTCGCTTCTCATGAGCGGCTTTGAGACAGTACACCAACCATGAGAGCTTCCACC 379

Db 485 CCGGCTTCTCATGAGCGGCTTTGAGACAGTACAGCCACATATGAGAGACTTTCATC 544

QY 380 GTAAGTATACAAATCCGCGCGACATGTACAGCTGCATCTGTGATCTCTTGCA 439

Db 545 GCAAGTGTATCAATCAATCCGCGCGACATGTACAGCTGCATCTGTGATCTCTTGCA 604

QY 440 ACCAAGCTTCTCCGCGCGACATGAGCGGCTGTATCTCATCTGAGGAGATGCTTCATCCG 499

Db 605 ACACACCTTCTCCGCGCGACATGAGCGGCTGTATCTCATCTGAGGAGATGCTTCATCCG 664

QY 500 TGTTCAGCTGTGATACCGGAGTCTTCATGATGAGTCAAGCTTCAGAGAGATCC 559

Db 665 TGTTCAGCTGTGATACCGGAGTCTTCATGATGAGTCAAGCTTCAGAGAGATCC 724

QY 560 TGAAGTCAATCTCTGCTGAAGAACAAGACCAAGAGAGCGGAGCTGCCATGTCA 619

Db 725 TGAAGTCAATCTCTGCTGAAGAATAAACAAGAGAGCGAGAGAGCTGCCATGTCA 784

QY 620 TCTGTGGAACAAGACCAAGCGGAGCTGTGCGGAGAGTCCCAACCAAGAGCGG 679

Db 785 TCTGTGGAACAAGACCAAGCGGAGCTGTGCGGAGAGTCCCAACCAAGAGCGG 844

QY 680 AGCTGCTGTGTGCGGAGAGAGTCCGCTTCTGAGTGTGCGGAGAGAGAGCA 739

Db 845 AGCTGCTGTGTGCGGAGAGAGTCCGCTTCTGAGTGTGCGGAGAGAGAGCA 904

QY 740 CCAAGTGAACGAATGTTCTAGCTGCTTTCAGCATGCGCAAGCTGCAACAGATGA 799

Db 905 CTAATGTGAACGAATGTTCTAGCTGCTTTCAGCATGCGCAAGCTGCAACAGATGA 964

QY 800 GCGCGCGGCGCATGCGAGAGTCTCCGCGAGTACAGTCCCTCCAGCGGCGCT 859

Db 965 GCGCGCGGCGCATGCGAGAGTCTCCGCGAGTACAGTCCCTCCAGCGGCGCT 1024

QY 860 TCTGCATGCGCGCGGTCAAGAGATGAGAGCTTATGAGATGATGCTTCCGCGCGC 919

Db 1025 TCTGCATGCGCGCGGTCAAGAGATGAGAGCTTATGAGATGATGCTTCCGCGCGC 1084

QY 920 GCGCGAGGTCAACAGTACCTCAAGTACATCAAGCGCAAGTCTTCCGAGAGCGCAG 979

Db 1085 GCGCGAGGTCAACAGTACCTCAAGTACATCAAGCGCAAGTCTTCCGAGAGCGCAG 1144

QY 980 CCGGTGAGAGGAGCAAGTGCATCTCAAGTGA---GCGAGGAGTCTGAGGCGGCGCTTG 1036

Db 1145 CCGGAGAGGAGCAAGTGCATCTCAAGTGAAGGAGGAGCATCGGAGAGGCGGCTTG 1204

QY 1037 GCGAGTGCCTTCAGGAGAGTGGCCAGATGCCCATGTGTGCGCATCTCC 1086

Db 1205 GCGAGTGCCTTCAGGAGAGTGGCTGAGAGATTCACATGCGCACATCCAC 1254

RESULT 12

ACH42415

ID ACH42415 standard; cDNA; 567 BP.

XX

AC ACH42415;

XX

DT 13-OCT-2003 (first entry)

XX

DE Human foetal brain cDNA #3782.

XX

KW Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;

KW genome mapping; biodiversity; genetic disorder.

OS Homo sapiens.

XX

PN US2003073623-A1.

XX

PD 17-APR-2003.

XX

PF 30-JUL-2001; 2001US-00918995.

PR 30-JUL-2001; 2001US-00918995.

PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.

PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

DR WPI; 2003-615964/58.

PT New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.

PS Claim 1; SEQ ID NO 29627; 44pp; English

The invention relates to an isolated polynucleotide comprising any one of the 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was determined by the technique of SBH (sequencing by hybridisation). Also included is a purified polypeptide comprising a sequence corresponding to a reading frame of the novel polynucleotide. The nucleic acid sequences are useful in diagnostics as expressed sequence tags (EST) for identifying expressed genes or for physical mapping of the human genome, in forensics, in assessing biodiversity, or in identifying mutations responsible for genetic disorders and other traits. The nucleotide sequences are also useful as hybridisation probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA. The purified polypeptide is useful for generating antibodies specific for it. The present sequence is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at [www.uspto.gov/sequence.htm?DocID=20030073623](http://www.uspto.gov/sequence.htm?DocID=20030073623)

SQ Sequence 567 BP; 138 A; 134 C; 147 G; 146 T; 0 U; 2 Other;

Query Match	16.1%	Score 496;	DB 8;	Length 567;
-------------	-------	------------	-------	-------------

Matches 499; Conservative 0; Mismatches 6; Indels 0; Gaps 0.

OY	2219	GGGAGGCGGGAGGAGGACGCTGGAGCCGAACTGAGCTGGGAGGGATCCGACAGAAA	227
Db	44	GGGAGCGGGAGGAGGACACTGGAGCCGAAACTGAGCTGGGAGGGATCCGACAGAAA	103
OY	2279	GCTCAGGGCGGGCTCTCTCTGTGTGCCGGGATTTGGGCTATGCTGGGTACCACTGTAC	233
Db	104	GCTCAGGGCGGGCTCTCTCTGTGTGCCGGGATTTGGGCTATGCTGGGTACCACTGTAC	163
OY	2339	TCAGGCATGTGGGTTTTGAACCCATMAACAAAGCCCTTGTCACTAGCTTTACAG	239
Db	164	TCAGGCATGTGGGTTTTGAACCCATMAACAAAGCCCTTGTCACTAGCTTTACAG	223
OY	2399	TATATTTGTATTTTATCTCTCTAAACATTTGAAGTTTGGGCCCTTAAGAACCTTA	245
Db	224	TATATTTGTATTTTATCTCTCTAAACATTTGAAGTTTGGGCCCTTAAGAACCTTA	283
OY	2459	GTGATCTTCTATTTGGGCTTTTCTGAGGTTCAAGAGGGTAATACCTTCTCAGGTAC	251
Db	284	GTGATCTTCTATTTGGGCTTTTCTGAGGTTCAAGAGGGTAATACCTTCTCAGGTAC	343
OY	2519	ACAGCAAGCTGTGGGTGSCAAGAGCAACTGACGCTGGGCACTTCACTATACCAAGAT	257
Db	344	ACAGCAAGCTGTGGGTGSCAAGAGCAACTGACGCTGGGCACTTCACTATACCAAGAT	403
OY	2579	GTGCTCCCTCTTGTAGCTTTGAGCCCTTGGGGCTTTCAGAGGCTTTGGGACATCTGTCT	263
Db	404	GTGCTCCCTCTTGTAGCTTTGAGCCCTTGGGGCTTTCAGAGGCTTTGGGACATCTGTCT	463
OY	2639	CAACCTCTCCCTAATCAGTCTGTGAGGGTCCCTGTAGATATTTGTACACCATGCCA	269

Db 464 CAACCTCTCCCTAGATCAATCTGTGAGGGTCCCTGTAGATATTGTGTACACCATGCCA 523

QY 2699 TGTATATACAAGTACACACAGATGT 2723

Db 524 TGTATATACANGTACACAGATGT 548

### RESULT 13

ID	ACD93776	standard; cDNA; 607 BP.
1	1	1

AC ACD93776;

DT 23-SEP-2003 (first entry)

DE Human colon cancer cell expressed cDNA #2188  
.....

Open reading frame detection; genome sequencing; colon cancer; breast cancer; population genome analysis; genetic shift; cancer; antibiotic resistance; antibiotic non-tolerance; congenital disease; agriculture; food crop genome; resistance gene; retrovirus; influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium; gene; ss.

OS Homo sapiens.

PN US2002155438-A1.

PD 24-OCT-2002.

PF 27-SEP-1999; 99US-00406117.

PR 20-NOV-1998; 98US-00196716.

PA (SIMP/) SIMPSON A J G.

PA (BREN/) BRENTANI R R.

PI Simpson AJG, Neto ED

DR WPI; 2003-182626/18.

PT Determining open reas

PT Determining open reading frames of genome of an organism e.g. a human  
 PT suffering from cancer involves use of single oligonucleotide primer and  
 PT low stringency for preparing single-stranded cDNA from mRNA of  
 PT individual.

PS Example 9; Page 336; 959pp; English.

CC The invention describes a method of detecting open reading frames in  
CC the genome of an organism, comprising contacting mRNA from cell of organism  
CC with a single oligonucleotide primer (1) at low stringency, preparing  
CC single-stranded cDNA by reverse transcribing mRNA with (1), amplifying  
CC cDNA, sequencing the product, and repeating the contacting, preparing  
CC and amplifying steps with different primers and sequencing resulting  
CC nucleic acids. The method is useful for: determining that a known  
CC nucleotide sequence from a genome of an organism corresponds to a  
CC nucleotide sequence of an open reading frame; for preparing a contig,  
CC nucleic acid molecule from a genome of an organism; and for sequencing  
CC all or part of a genome of an organism. mRNA is obtained from mammalian  
CC or human cell which is associated with a pathological condition e.g. a  
CC colon cancer or breast cancer cell. The method is useful for analyses of  
CC populations of subjects and can be used to carry out genetic analyses of  
CC large or small populations. further, it can be used to study living  
CC systems to determine if, e.g. there have been genetic shifts which render  
CC an individual or population more or less likely to be afflicted with  
CC diseases such as cancer, to determine antibiotic resistance or non-  
CC tolerance, and so forth. The method can also be used in the study of  
CC congenital diseases, and the risk of affliction to a foetus, as well as  
CC the study of whether the conditions are likely to be passed to offspring  
CC through ova or sperm. The analyses for pathological conditions can be  
CC carried out in all animals, plants, birds, fish, etc. Using this method,  
CC in the area of agriculture, for example the genomes of food crops can be





Db 402 CTGGTACCACTGATCTGAGCATGCTGGTTTGAACCAACCAAGCCCTTG 461  
 QY 2381 TCATCAG 2387  
 Db 462 TCATCAG 468

RESULT 15  
 ID ACC42325  
 ACC42325 standard; cDNA; 1744 BP.

ACCA2325;

22-MAY-2003 (first entry)

Human MAP kinase cascade activator #44 cDNA.

Human; ELK1 phosphorylation; ELK1 phosphorylation kinase; vinticide;  
 antiinflammatory; immunomodulator; cyostatic; anti-allergic; anti-HIV;  
 antiinflammatory; antiarthritic; antidiabetic; antiaesthetic; gene therapy;  
 inflammation; autoimmune disease; viral disease; cancer; diabetes;  
 rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;  
 Iga nephritis; gene; ss.

Homo sapiens.

WO2003008589-A1.

30-JAN-2003.

15-JUL-2002; 2002WO-JP007174.

18-JUL-2001; 2001JP-00218204.

31-AUG-2001; 2001JP-00263450.

21-JAN-2002; 2002JP-00012176.

(ASAH) ASAHIT KASEI KOGYO KK.

Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;

WPI: 2003-229582/22.

P-PSDB; ABR41059.

Elk1 phosphorylation-associated gene and its encoded protein with MAP

kinase cascade effect, applicable in diagnosis of and developing drugs

for e.g. inflammations, autoimmune diseases, viral diseases and cancer.

Claim 4; Page 316-320; 762pp; Japanese.

The invention relates to a novel purified protein having ELK1

phosphorylation activity and/or an activity of activating ELK1

phosphorylation kinase. A protein of the invention has antiinflammatory,

immunomodulator, vinticide, cyostatic, antiallergic, antiaesthetic,

antiarthritic, antidiabetic, antiaesthetic, and anti-HIV activity. The

polynucleotides may have a use in gene therapy. The gene and its encoded

protein are applicable in diagnosis of and developing drugs for e.g.

inflammations, autoimmune diseases, viral diseases and cancer such as

rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral

hepatitis and Iga nephritis. The present sequence is used in the

exemplification of the invention

Sequence 1744 BP; 372 A; 575 C; 503 G; 294 T; 0 U; 0 Other;

Query Match 13.5%; Score 416; DB 7; Length 1744;

Best Local Similarity 65.9%; Pred. No. 8.4e-53;

Matches 668; Conservative 0; Mismatches 310; Indels 36; Gaps 3;

QY 25 AGCCGCGCCAGCCGCGCTCCGAGCAGCGCAGGAGAGTCCCGCGAGTACCCG 84

Db 44 AGCCGCGCTTACGCGAGCCAGCGAGCGAGCGAGCCCGAGCCCGAGCCCGCC 103

QY 85 GAGCCACACACAGCTCTGGAGGCTCGCGGCTGGAGCAGCAGCAGCTCCCGCAGCTC 144

Db 104 CAGCCGAGAGAGCCCTCCAGCGCTCACCCCGCGCCAGCCCGCAGCCCTCAGCCG 163  
 QY 145 CCGCCGCTTCCAGGACCTCTGAGCGCGCCAGAGCCCGCCATTCAGCCG 204  
 Db 164 CTCTCTCCCTTCTCTCGCGCCCGCCGCGCCCTTCCGCGCCCTTCCGCAATGAACTG 223  
 QY 205 CGAGCCATGATGAAAGCTTTGTCAGCGGAGCTGACGCTCAGTGGCCCGCAAAAC 264  
 Db 224 GCGCGATGATCAAGAGATGTGCGCGAGCAGCTCGAGTGTATCCCGGCAAGAAC 283  
 QY 265 TCATACCGCATGT 324  
 Db 284 TGCTATGATGT 343  
 QY 325 TTCTCTAATGCGCGCTTGTGAGACCAATACACACCACTTCAAGAGCTTCCACCGTAAG 384  
 Db 344 TTCTCTACCGCGCGCTTGTGAGAGCGCTTACACCGCTTCAATCAGAGACTTCCACCGTAAG 403  
 QY 385 GTATACAAACATCGCGCGGAGCATGTACAGCTGACATCTGTGATCCTTGTGCAACAC 444  
 Db 404 TTCTATCTCATTCGCGCGAGGTCTTACAGCTGACATCTTCCAGAGCTTGTGTGTGTGT 463  
 QY 445 CCCTTCCCGCATGCGCAGGCTGTCTCATCTTCAAGGAGATGTCTTCTGTGTGTGTGT 504  
 Db 464 CCGTTCGCGCATGCGCGCGCTCTCATCTTCAAGGAGATGTCTTCTGTGTGTGTGT 523  
 QY 505 AGCTGTGATTAACCGGAGGTCTTGTGATGAGGTCAAGGCGCTTCAAGAGCATCTTGTGAG 564  
 Db 524 AGTCTGACAAACCGGAGGTCTTGTGAGAGGTGAGGCGCTTCAAGAGCATCTTGTGAG 583  
 QY 565 GTCAAGTCTGCTGAGAAACAAAGCAAGAGAGGCGCGAGCTGCCATGTGTGTGTGTGT 624  
 Db 584 ACCAGTCTTGTCTTCAAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643  
 QY 625 GGCACAAAG 684  
 Db 644 GGCACAAAG 700  
 QY 685 CTGCTGT 741  
 Db 701 CTGCTGT 760  
 QY 742 AACGTGAG 801  
 Db 761 AGCTGTGAG 820  
 QY 802 CCGCGCTTGT 857  
 Db 821 CCGAGCTGT 880  
 QY 858 -----CTTGTGATGCGCGCGCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 891  
 Db 881 CCGAG 940  
 QY 892 TATGCGATGT 951  
 Db 941 TTTGCGATGT 1000  
 QY 952 AAGGCAAG 1005  
 Db 1001 CGCGAG 1054

Search completed: June 20, 2004, 03:15:33  
 Job time: 1146 secs



PF 28-JAN-2002; 2002MO-US002301.  
XX  
PR 08-FEB-2001; 2001US-00778963.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Neelam B, Ketchum KA, Di Francesco V, Beasley EM,  
XX  
XX WPI: 2002-706901/76.  
DR P-PSDB; ABG70951.  
DR

PT New human Ras-like polypeptides, useful for treating diseases associated  
with an increase in apoptosis or cell proliferation e.g., cancer.  
XX  
Claim 4; Fig 1A; 73pp; English.  
PS

QY	12	SerGlyAspLeuYAsnAspThrIysGluAlaIaGluLeuPrometValIleCysGlyAsn	140
Db	571	TCTCGCTGAAGAACAAAGACCAAGAGAGCGCGGAGCTGCCATGTCATCTGTGGCAAC	630
QY	141	LysAsnAspHisGlyGluLeuGlyAspArgGlnValProThrArgIuaIaGluLeuLeuVal	160
Db	631	AAGAACCAACACGGCGAGCTGTGCCGCCAGGCGCCACCAACGAGCGAGCTGTGTG	690
QY	161	SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIysIysAsnThrAsnValAsp	180
Db	691	TCGGGCGACGAGAACTCCGCTCACTTCGAGGTGTGGCGCAAGAGAAACCAACAGTGGAC	750
QY	181	GluMetPheTyrValLeuPheSerMetAlaIysLeuProHisGluMetSerProAlaLeu	200
Db	751	GAAATGTTCACCTGCTTTCAGCATGGCCAAAGCTGCACACAGAGATGAGGCCCGGCTGT	810

The present invention relates to a new polypeptide comprising a fully defined 266-amino acid sequence or its fragment, orthologue or allelic variant. The allelic variant or orthologue is encoded by a nucleic acid molecule that hybridises under stringent conditions to the opposite strand of a nucleic acid molecule with a fully defined sequence of 3082 or 1121 bp. The pharmaceutical composition comprising the agent that binds to the polypeptide is useful for treating a disease or condition mediated by a human Ras-like protein e.g. cancer. Likewise, the polypeptide is useful for identifying modulators of Ras-like protein activity or expression by screening a compound for the ability to stimulate or inhibit interaction between the Ras-like protein and an interacting molecule or Ras-like activity. The polypeptide is useful for diagnosing a disease, or predisposition to a disease, or treating diseases associated with an increase in apoptosis or cell proliferation e.g. AIDS (acquired immunodeficiency syndrome), Alzheimer's disease, Parkinson's disease, aplastic anaemia, allergies, atherosclerosis, osteoporosis or bacterial, viral, fungal, parasitic, protozoal or helminthic infections, and inflammation. The present nucleic acid sequence encodes the human Ras-like protein of the invention. This gene is located on chromosome 22

Sequence 3082 BP; 708 A; 962 C; 793 G; 619 T; 0 U; 0 Other;

Predictive Accuracy:	
Pred. No.:	3,61e-148
Score:	1377.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
	Gaps: 0
	Length: 3082
	Matches: 266
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-09-778-963A-2 (1-266) X ABS54600 (1-3082)

[illegible]

QY	241	AsnSerAspLeuIlysrVrIleYsAlaYsValleuIrgIugIyGlnAlaRGluar	260
Db	931	AAcAGTGAcCTCAAGACATCAAGGCAAGGTCTCTGGGAAGGCGCAGGCCGTGAGAGG	990
OY	261	AspIlyeCYeThrIleGln	266
Db	991	GACAACTGCACCATTCAG	1008

RESULT 2  
AAZ36913 standard, DNA, 837 BP.

DT 13-MAR-2000 (first entry)

DE DNA encoding a homologue of G protein signalling AGS1.

KM Activator of G protein signalling; AGS; ras-related G protein;  
KM GTP hydrolysis; G protein activity; pheromone response pathway  
KM G protein-coupled signal transduction; G-gamma selectivity;  
KM cellular signal transduction; AGS1 homologue; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

/product= "AGS1 homologue"

PN W09958670-A1.

PD 18-NOV-1999  
XX

0 / - MAY - 1999; 99WC-US010151.  
PF  
XX

PR 07-OCT-1998: 98US-0103355P

PA (CADU-) CADUS PHARM CORP.

PI Cismowski M, Duzic E;

DR WPI; 2000-072337/06.

2000

PT characterized by an aberrant AGS protein activity.

PS Example 15; Page 143-144; 162pp; English.

XX The present sequence encodes a homologue of a human AGS1 (activator of G protein signaling (AGS1) protein. The AGS1 cDNA sequence was isolated from a human liver cDNA library. The AGS1 protein exhibits homology to ras-related G proteins, and contains alterations in conserved amino acids consistent with a deficiency in GTP hydrolysis activity. AGS1 stimulates G protein activity, G protein-coupled signal transduction and the pheromone response pathway in a receptor-independent manner. The AGS1 protein also shows G-gamma selectivity, as measured by growth assays in yeast. CC expressing various mammalian G-gamma constructs, and tissue-specific expression, as measured by Northern blot analysis. The AGS1 protein can be used to screen for compounds that modulate cellular signal transduction. CC The protein is used to treat disorders characterized by an aberrant AGS1 protein activity or AGS1 nucleic acid expression

XX Sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	2,24e-148	Length:	837
Score:	1372.00	Matches:	265
Percent Similarity:	99.62%	Conservative:	0
Best Local Similarity:	99.62%	Mismatches:	1
Query Match:	99.64%	Indels:	0
DB:	3	Gaps:	0

US-09-778-963a-2 (1-266) x AAZ36913 (1-837)

QY 1 MetMetlyThrIleuSerSerGlyYasnCyThrIleuSerValProAlaIysAsnSerTy 20  
 DB 37 ATGATGAAGACCTTGTCCAGCGGGAACGACCGCTCAGTGGCCGCCAAAGCATATC 96  
 QY 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerIleu 40  
 DB 97 CGCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 156  
 QY 41 AsnGlyArgPheGluAspGlnTyThrProThrIleGluAspPheIleArgIleValTy 60  
 DB 157 AATGGCCGCTTGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 216  
 QY 61 AsnIleArgGlyAspMetTyGlnIleuAspIleuAspThrSerGlyAsnIleProPhe 80  
 DB 217 AACATCCGGGAGCATGTACAGCTCGACATCCGATACCTCTGGACACACCCCTTC 276  
 QY 81 ProAlaMetArgIleuSerSerIleuThrGlyAspValPheIleuValPheSerIleu 100  
 DB 277 CCCGCGATGCGGAGCTGTCCATCTCCACAGGGAGTGTCTTCATCTGTTGACGCTG 336  
 QY 101 AspAsnArgGlyIleuSerPheAspGluValIysArgIleuGlnIleuGluValIys 120  
 DB 337 GATTAACCGGAGTCTTGTGATGAGGTCAAGCGCTTCAGAAAGCATCTCGAGGTCAAG 396  
 QY 121 SerCyIleuIysAsnIleuThrIleuGlnIleuAspIleuAspProMetValIleuCy 140  
 DB 397 TCCCTCCCGAAGAACAAACCAAGAGGCGGAGGCTGCCATGTCTATCTGTGGCAAC 456  
 QY 141 LysAsnAspPheIleuGlnIleuCyArgIleuValProThrThrGlnIleuGlnIleuVal 160  
 DB 457 AAGAAACGACCAAGCGGAGCTGTGGCCAGGTCGCCACACCGAGCGGAGTGTGGTG 516  
 QY 161 SerGlyAspGluAsnSerIleuValPheGluIleuSerAlaIysIleuAsnThrAsnVal 180  
 DB 517 TCGGGCGAGCAAGAACGCGCTTCACTTCAGAGGTGTGGCCAAAGAACACCAACGTGAC 576  
 QY 181 GluMetPheTyValIleuPheSerMetAlaIysIleuProIleuGlnIleuSerProAla 200  
 DB 577 GAGATGTTCTACGTCTCTTACGATGGCCAAAGCTGCCACAGAAATAGACCCCGCTG 636  
 QY 201 HisArgIysIleSerValGlnTyGlyAspAlaPheIleProArgProPheCyMetArg 220  
 DB 637 CATCCAGAGATCTCGTGCAGTACGAGTACGCTTCCACCCAGGCGCTTGTGATGGC 696  
 QY 221 ArgValIysGluMetAspAlaTyGlyMetValSerProPheAlaArgIleProSerVal 240

DB 697 CGCGTCAAGGAGATGACGCGCTTATGATGCTGTGCGCTTCCGCCCGCCCGCCAGCGTC 756

QY 241 AsnSerIleuIysTyThrIleuValIleuArgIleuGlnIleuValIleuValIleuVal 260

DB 757 AACGTAACCTCAAGTACATCAAGGCAAGGTCTTGGAGAGCGCCGCGTGAAGG 816

QY 261 AspIysCyThrIleGln 266

DB 817 GACAAATGACCATTCAG 834

## RESULT 3

ABL92076

ID ABL92076 standard; cDNA; 2832 BP.

AC ABL92076;

DT 30-MAY-2002 (first entry)

DE Human Tumour Endothelial Marker polynucleotide SEQ ID NO 174.

XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;

XX normal endothelial marker; pan-endothelial marker; immunostimulant;

XX antiangiogenic; tumour; neovascularization; vascularised tumour;

XX polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;

XX psoriasis; gene; ss.

OS Homo sapiens.

PN WO200210217-A2.

PD 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US024031.

XX 02-AUG-2000; 2000US-0222599P.

XX 11-AUG-2000; 2000US-0224360P.

XX 11-APR-2001; 2001US-0282850P.

XX (UWUO) UNIV JOHNS HOPKINS.

PI St Croix B, Kinzler KW, Vogelstein B;

XX WPI; 2002-291856/33.

DR An isolated molecule comprising an antibody variable region which

XX specifically binds to an extracellular domain of a tumor endothelial

XX marker (TEM) protein, useful for inhibiting tumor growth.

XX Disclosure; Page 119-120; 33pp; English.

PS The invention relates to an isolated molecule comprising an antibody

XX variable region which specifically binds to an extracellular domain of a

XX tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,

XX ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM

XX proteins have cytostatic, immunostimulant and antiangiogenic activity.

XX They are useful for inhibiting tumour growth, neovascularization in subjects

XX bearing a vascularised tumour, polycystic kidney disease, diabetic

XX retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM

XX genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90769)

XX are disclosed, as are marker oligonucleotide sequences: tumour

XX endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191, normal

XX endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers

XX (PEM) ABL91903-ABL91995

XX Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1,21e-147	Length:	2832
Score:	1372.00 <td>Matches:</td> <td>265</td>	Matches:	265
Percent Similarity:	99.62% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.62% <td>Mismatches:</td> <td>1</td>	Mismatches:	1
Query Match:	99.64% <td>Indels:</td> <td>0</td>	Indels:	0

```

DB: 6 Gaps: 0
US-09-778-963a-2 (1-266) x ABX72001 (1-2832)
QY 1 MetMetLysThrLeuSerSerGlyAsnCyThrLeuSerValProAlaLysAsnSerTyr 20
DB 37 ATGATGAAGACTTTGTCAGCGGGAAGTGCACGCTCAGTGTCCCGCCAAAATCATAC 96
QY 21 ArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40
DB 97 CGCATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60
DB 157 AATGCGCGCTTTGAGAGACCAAGTACACCCACCATCGAGACTTCCACCGTAAAGTATAC 216
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
DB 217 AACATCCGCGGCGCATGTACACGCTCCGATCTCGAATCCTCTGGCAACCAACCCCTTC 276
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
DB 277 CCGGCGCATGCGAGGCTGTCTCCTCAGGAGATGCTCTTCTGCTGCTGCTGCTGCTGCTG 336
QY 101 AspaAsnArgGluSerPheAspGluValLysArgLeuGlnIleLeuGluValLys 120
DB 337 GATTAACCGGAGCTCTTGCATGAGGTCAAGCCCTTCAGAACCAAGTCTTGAGAGTCAAG 396
QY 121 SerCySLeuLysAsnLysThrLysGluAlaAlaGluLeuProMetValIleCySGLys 140
DB 397 TCCTCCCTGAGAACACAGACCAAGAGGCGGAGCTGCCATGCTCATCTGTGGCAAC 456
QY 141 LysAsnAspHisGlyGluLeuGlyArgGlnValProThrThrGluAlaGluLeuVal 160
DB 457 AAGAACGACCAAGCGAGCTGTGCGCCAGGCGCCACCAAGCGCTGTGCTG 516
QY 161 SerGlyAspGluAsnSerIleArgPheGluValSerAlaLysAsnThrAsnValAsp 180
DB 517 TCGGCGAGAGAACTGGGCTTCTTCCAGGTGTGCGCCCAAGAACCAACCACTGAC 576
QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
DB 577 GAGATGTTCTAGCTCTTTCAGCATGGCCAGCTGCCACAGAGTAGAGCCCGCCCTG 636
QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCySLeuArg 220
DB 637 CATCCCAAGATCTCCGTGAGTACGAGTACGCTTCCACCCAGGCGCTTCTGATGGCC 696
QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240
DB 697 CGCGTCAAGAGATGAGCCCTTATGGCATGTCTGCGCTTGGCCCGCCCGCCAGCTTC 756
QY 241 AsnSerAspLeuLysTyrIleLysAlaLysValLeuArgGluGlnAlaArgGluArg 260
DB 757 AACAGTGAACCTCAAGTACATCAAGGCAAGGCTCTTCCGGAAGGCGCCGCTGAGAG 816
QY 261 AspLysCyThrIleGln 266
DB 817 GACAGTGCACCATCCAG 834

```

```

KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
KW neovascularization; immune response; cytoskeletal; gene;
KW ophthalmological; antineoplastic; antirheumatic; antiarthritic; antipsoriatic; ds.
OS Homo sapiens.
PN WO200283874-A2.
XX 24-OCT-2002.
XX 10-APR-2002; 2002WO-US0008253.
XX 11-APR-2001; 2001US-0282850P.
XX 06-FEB-2002; 2002US-0354262P.
XX (U) UNIV JOHNS HOPKINS.
XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;
XX MPI: 2003-093016/08.
XX P-FSDB: ABUS4429.
XX New purified human transmembrane protein, designated as tumor endothelial
XX marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
XX polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
XX psoriasis.
XX Disclosure: Page 113-114; 374pp; English.
XX The present invention relates to a novel method for the isolation of
XX endothelial cells (ECs), and the identification of genes expressed in
XX normal and tumor ECs. Tumour endothelial marker (TEM), normal
XX endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
XX identified in human ECs. The human EC marker proteins and the
XX polynucleotide sequences encoding them are useful for detecting,
XX diagnosing or treating tumors as well as polycystic kidney disease,
XX diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
XX useful for inhibiting neovascularization or tumour angiogenesis, for
XX inducing an immune response to tumour endothelial cells in a patient, or
XX for identifying candidate drugs for treating tumors. The present
XX sequence represents a human TEM or NEM gene of the invention
XX
SQ Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 1,216-147	2832	265	0	1	0	0
Percent Similarity: 99.62%						
Best Local Similarity: 99.62%						
Query Match: 99.64%						

```

US-09-778-963a-2 (1-266) x ABX72001 (1-2832)
QY 1 MetMetLysThrLeuSerSerGlyAsnCyThrLeuSerValProAlaLysAsnSerTyr 20
DB 37 ATGATGAAGACTTTGTCAGCGGGAAGTGCACGCTCAGTGTCCCGCCAAAATCATAC 96
QY 21 ArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40
DB 97 CGCATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60
DB 157 AATGCGCGCTTTGAGAGACCAAGTACACCCACCATCGAGACTTCCACCGTAAAGTATAC 216
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
DB 217 AACATCCGCGGCGCATGTACACGCTCCGATCTCGAATCCTCTGGCAACCAACCCCTTC 276
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
DB 277 CCGGCGCATGCGAGGCTGTCTCCTCAGGAGATGCTCTTCTGCTGCTGCTGCTGCTGCTG 336

```

QY 101 AspaanArgIuSerPheaspGluValLysArgLeuGlnIleLeuGluValLys 120  
 DB 337 GATTAACGGAGGCTCTTCATGATGAGTCAAGGCCCTTCGAAAGCATCTCGAGGTCAAG 396  
 QY 121 SerCysLeuLysAsnLysThrLysGluAlaIaGluLeuPrometValIleCysGlyAsn 140  
 DB 397 TCCTGCTCGAAGAACCAAGACCAAGAGGCGCGAGCTGCCCATGCTCATCTGTGGCAAC 456  
 QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160  
 DB 457 AAGAACGACCAACGGGAGCTGTGCGCCAGGCGCCACACCGAGCCGAGCTGTGTG 516  
 QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180  
 DB 517 TCGGCGCAGAGAACCTGCGCTTCACTTCAGAGGTGTGCGCAAGAAACACCAACCTGTGAC 576  
 QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200  
 DB 577 GAGATGTTCTACGTCTCTTCAGCATGGCCCAAGCTGCCACACGAGATGAGCCCGCCCTG 636  
 QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
 DB 637 CATCGCAAGATCTCGTGCAGTACGAGTACCGCTTCCACCCAGGCGCTTCTGATGGCG 696  
 QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheHisArgArgProSerVal 240  
 DB 697 CGGTCAGAGGATGAGCGCTTATGGCAATGCTCTGCGCTTCCGCGCGCCCGCCGCTG 756  
 QY 241 AsnSerAspLeuLysTyrIleLysAlaLysValLeuArgGlnGlyAlaArgGluArg 260  
 DB 757 AACAGTACCTCAAGTACATCAAGGCCCAAGGCTCTTCCGGAAGGCCAGGCCCGTGAAGG 816  
 QY 261 AspLysCysThrIleGln 266  
 DB 817 GACCAAGTCACCATCCAG 834

## RESULT 5

ABL92087  
 ID ABL92087 standard; cDNA; 2973 BP.

AC ABL92087;

DT 30-MAY-2002 (first entry)

DE Human Tumour Endothelial Marker polynucleotide SEQ ID NO 197.

XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostratic;

KW normal endothelial marker; pan-endothelial marker; immunostimulant;

KW antiangiogenic; tumour; neovascularisation; vasculature; tumour;

KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;

XX psoriasis; gene; ss.

XX Homo sapiens.

OS WO200210217-A2.

PD 07-FEB-2002.

PF 01-AUG-2001; 2001WO-US024031.

PR 02-AUG-2000; 2000US-0222599P.

PR 11-AUG-2000; 2000US-0224360P.

PR 11-APR-2001; 2001US-0282850P.

XX (UWJO ) UNIV JOHNS HOPKINS.

XX St Croix B, Kinzler KW, Vogelstein B;  
 XX WPI; 2002-291856/33.  
 DR P-PSDB; ABB90733.  
 XX An isolated molecule comprising an antibody variable region which

PT specifically binds to an extracellular domain of a tumor endothelial  
 marker (TEM) protein, useful for inhibiting tumor growth.

PS Claim 65; Page 152-153; 33pp; English.

CC The invention relates to an isolated molecule comprising an antibody  
 CC variable region which specifically binds to an extracellular domain of a  
 CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,  
 CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM

CC proteins have cytostratic, immunostimulant and antiangiogenic activity.  
 CC They are useful for inhibiting tumour growth, neovascularisation in subjects  
 CC bearing a vascularised tumour, polycystic kidney disease, diabetic

CC retnopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM  
 CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)

CC are disclosed, as are marker oligonucleotide sequences: tumour  
 CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191, normal

CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers  
 CC (PEM) ABL91903-ABL91995

CC Sequence 2973 BP; 596 A; 967 C; 787 G; 623 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1,292-147 Length: 2973  
 Score: 1372.00 Matches: 265  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.64% Indels: 0  
 DB: 6 Gaps: 0

US-09-778-963A-2 (1-266) x ABL92087 (1-2973)

QY 1 MetMetLysThrLeuSerSerGlyAsnCysThrIleuSerValProAlaLysAsnSerTyr 20

DB 208 ATGATGAAGACTTGTCTCAGCGGAGACCTGCAAGTGTGCGCCGCAAAAGCTCATAC 267

QY 21 ArgMetValIleLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40

DB 268 CGCATGGTGTGCTGGGTGCTCTCGGGTGGCAAGAGCTCATCGTGTGCTTCTC 327

QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60

DB 328 AATGGCGCTTGTGAGGACACAGTACACACCCATCAGAGACTTCCACCGTAAGGTATAC 387

QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80

DB 388 AACATCCGCGCGACATGACACACTGCACATCTGTGATCTGTGCAACACCCCTTC 447

QY 81 ProAlaMetArgAlaGluSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100

DB 448 CCGCCATGCGCAGAGCTGTGCTCATCTTCACAGGAGATGCTTCACTGCTTCAAGCTTC 507

QY 101 AspaanArgIuSerPheaspGluValLysArgLeuGlnIleLeuGluValLys 120

DB 508 GATTAACGGAGGCTCTTCATGATGAGTCAAGGCCCTTCGAAAGCATCTCGAGGTCAAG 567

QY 121 SerCysLeuLysAsnLysThrLysGluAlaIaGluLeuPrometValIleCysGlyAsn 140

DB 568 TCCTGCTCGAAGAACCAAGACCAAGAGGCGCGAGCTGCCACAGGATCATCTGTGCAAC 627

QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160

DB 628 AAGAACGACCAACGGGAGCTGTGCGCCAGGCGCCACACCGAGCCGAGCTGTGTG 687

QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180

DB 688 TCGGCGCAGAGAACCTGCGCTTCACTTCAGAGGTGTGCGCAAGAAACACCAACCTGTGAC 747

QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200

DB 748 GAGATGTTCTACGTCTCTTCAGATGAGTGCACACGAGATGAGAGCCCGCCCTG 807

QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220



Db 808 CATGCAAGATCTCCGTGAGTACGATGACGCTTCCACCCAGCCGCTTCGATGCGC 867  
Qy 221 ArgVallylsGIuMetAspAlaTyrlGlyMetValSerProPheAlaArgProSerVal 240  
Db 868 CGCGCAAGAGATGACGCGCTATGCGATGCTCGCCCTTCGCCGCCGCCGCCGCCGCTC 927  
Qy 241 AsnSerAspLeuLysTyrlIleValAlaLysValIleuArgGluGlyAlaArgGluArg 260  
Db 928 AACACTGACCTCACTCACTCACTCAAGGCCAAGGCTCTTCGGGAAGGCCAGCCCTGAGAG 987  
Qy 261 AspLysCysThrIleGln 266  
Db 988 GACCAAGTGCACCATCCAG 1005  
RESULT 6  
ABX72012  
ID ABX72012 standard; DNA; 2973 BP.  
AC ABX72012;  
XX  
XX 12-MAR-2003 (first entry)  
DT  
XX  
XX DNA encoding human tumour endothelial marker TEM 2.  
DE  
XX  
XX Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
KW Tumour endothelial marker; normal endothelial marker; PEM;  
KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
KW neovascularization; immune response; cystostatic; antidiabetic; gene;  
KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200283874-A2.  
PN  
XX 24-OCT-2002.  
PD  
XX 10-APR-2002; 2002MO-US008253.  
PF  
XX 11-APR-2001; 2001US-0282850P.  
PR 06-FEB-2002; 2002US-0354262P.  
XX  
XX (UYUO ) UNIV JOHNS HOPKINS.  
PA  
XX  
XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;  
PI WPI; 2003-093016/08.  
DR P-PSDB; ABU54440.  
XX  
XX New purified human transmembrane protein, designated as tumor endothelial  
PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,  
PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
PT psoriasis.  
XX  
XX  
XX Disclosure; Page 154-155; 374pp; English.  
XX  
XX The present invention relates to a novel method for the isolation of  
CC endothelial cells (ECs), and the identification of genes expressed in  
CC normal and tumour ECs. Tumour endothelial marker (TEM), normal  
CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
CC identified in human ECs. The human EC marker proteins and the  
CC polynucleotide sequences encoding them are useful for detecting,  
CC diagnosing or treating tumors as well as polycystic kidney disease,  
CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for  
CC inducing an immune response to tumour endothelial cells in a patient, or  
CC for identifying candidate drugs for treating tumors. The present  
CC sequence represents a human TEM or NEM gene of the invention.  
XX  
XX Sequence 2973 BP, 596 A, 967 C, 787 G, 623 T, 0 U, 0 Other;  
SQ  
Alignment Scores: 1.29e-147 Length: 2973  
Pred. No.:

Score: 1372.00 Matches: 265  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.64% Indels: 0  
DB: 7 Gaps: 0  
US-09-778-963A-2 (1-266) x ABX72012 (1-2973)  
Qy 1 MetMetLysThrIleuSerSerGlyAsnCythrIleuSerValProAlaLysAsnSerTyr 20  
Db 208 ATGATGAAGACTTGTTCAGGGGAACTGACAGCTCACTGATGCGCCGCAAAACCTCAATAC 267  
Qy 21 ArgMetValIleuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheIleu 40  
Db 268 CGCATGGTGGTGTCTGGGTGCTCTCGGGTGGGCAAGACTTCATGCTCTCGCTTCCTC 327  
Qy 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheIleuArgValTyr 60  
Db 328 AATGGCGCTTTGAGAGACAGTACACACCCACATCGAGACTTCCACCGTAAGGTATAC 387  
Qy 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
Db 388 AACATCCGCGGACATGTATACACTGACATCTCTGTGATACCTTGGCAACCCCTTC 447  
Qy 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
Db 448 CCGGCCATGCCAGAGCTCTCATCTCTCAAGGGATGCTTCACTCTGCTTCACTGCTG 507  
Qy 101 AspAsnArgGluSerPheAspGluValIleuArgLeuGlnIleLeuGluValIle 120  
Db 508 GATTAACCGGAGTCTCTTCATGATGAGTCAAGGCGCTTCGAGAAGACATCTGAGGTCAAG 567  
Qy 121 SerCysLeuLysAsnLysThrIleGluAlaIleGluLeuProMetValIleCysGlyAsn 140  
Db 568 TCTGCTGAAAGAACAAAGACAGAGAGCGGAGAGCTGCCATGATGATGAGGAC 627  
Qy 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160  
Db 628 AAGAACAGACCAAGCGGAGCTGCGCGGAGTGCACCAAGCGGAGCTGCTG 687  
Qy 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180  
Db 688 TCGGCGCAGAGAACTGCGCTCACTTCAGAGTGTGCGGCAGAGAAACACCAAGCTGAC 747  
Qy 181 GluMetPheTyrValIleuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200  
Db 748 GAGAGTTCAGCTGCTCTTCAGCATGCGCAAGCTGCGCACAGAGATGAGCCCGCCTG 807  
Qy 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
Db 808 CATGCAAGATCTCGTGCATGATGAGTACCGCTTCCACCCAGGCCCTTTCGATGCGC 867  
Qy 221 ArgVallylsGIuMetAspAlaTyrlGlyMetValSerProPheAlaArgProSerVal 240  
Db 868 CGCGTCAAGAGATGACGCGCTATGCGATGCTTCGCCCTTCGCCGCCGCCGCCGCCGCTC 927  
Qy 241 AsnSerAspLeuLysTyrlIleValAlaLysValIleuArgGluGlyAlaArgGluArg 260  
Db 928 AACACTGACCTCACTCACTCAAGGCCAAGGCTCTTCGGGAAGGCCAGCCCTGAGAG 987  
Qy 261 AspLysCysThrIleGln 266  
Db 988 GACCAAGTGCACCATCCAG 1005  
RESULT 7  
ABK71563  
ID ABK71563 standard; cDNA; 3427 BP.  
AC ABK71563;  
XX  
XX 30-JUN-2002 (first entry)  
DT  
XX  
XX Human dithp polynucleotide #29.

XX Human; dthp; diagnostic and therapeutic polynucleotide; gene; ss; bone;  
 KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;  
 KW inflammatory disorder; viral infection; bacterial infection; seizure;  
 KW fungal infection; parasitic infections; developmental disorder; cervix;  
 KW endocrine disorder; metabolic disorder; neurological disorder; cervix;  
 KW gastrointestinal disorder; transport disorder; gene therapy; kidney;  
 KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;  
 KW skin; testis; thymus.  
 XX Homo sapiens.  
 OS WO200220754-A2.  
 XX 14-MAR-2002.  
 PD 29-AUG-2001; 2001WO-US027127.  
 XX 05-SEP-2000; 2000US-0229747P.  
 XX 05-SEP-2000; 2000US-0229748P.  
 PR 05-SEP-2000; 2000US-0229749P.  
 PR 05-SEP-2000; 2000US-0229750P.  
 PR 05-SEP-2000; 2000US-0229751P.  
 PR 05-SEP-2000; 2000US-0230583P.  
 PR 06-SEP-2000; 2000US-0230585P.  
 PR 06-SEP-2000; 2000US-0230514P.  
 PR 06-SEP-2000; 2000US-0230515P.  
 PR 06-SEP-2000; 2000US-0230517P.  
 PR 06-SEP-2000; 2000US-0230518P.  
 PR 06-SEP-2000; 2000US-0230519P.  
 PR 06-SEP-2000; 2000US-0230595P.  
 PR 06-SEP-2000; 2000US-0230597P.  
 PR 06-SEP-2000; 2000US-0230598P.  
 PR 06-SEP-2000; 2000US-0230599P.  
 PR 06-SEP-2000; 2000US-0230610P.  
 PR 06-SEP-2000; 2000US-0230865P.  
 PR 07-SEP-2000; 2000US-0230988P.  
 PR 07-SEP-2000; 2000US-0230951P.  
 PR 07-SEP-2000; 2000US-0231163P.  
 PR 07-SEP-2000; 2000US-0231167P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Stuart J, Lincoln SE, Alcus CM, Dufour GE, Chalup MS, Hillman JL;  
 PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;  
 PI Moniyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;  
 PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;  
 PI Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR;  
 XX WPI; 2002-383054/41.  
 DR P-PSDB; ABG59971.  
 XX  
 PT An isolated polynucleotide useful in diagnostics and therapeutics.  
 PS Claim 1; Page 419-420; 686pp; English.  
 XX  
 CC The invention relates to human diagnostic and therapeutic (dthp)  
 CC polynucleotides and their associated polypeptides (DTHP polypeptides).  
 CC The sequences of the invention are used in the treatment and diagnosis of  
 CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers  
 CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,  
 CC cervix, kidney, lung, ovary, pancreas, prostate, spleen, testis or  
 CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,  
 CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal  
 CC infections, parasitic infections, developmental disorders (e.g. anaemia,  
 CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),  
 CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders  
 CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,  
 CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal  
 CC disorders (e.g. ulcerative colitis, lymphoma) and transport disorders  
 CC (e.g. myotonic dystrophy, catatonia, peripheral neuropathy). Sequences  
 CC ABK71535-ABK71809 represent human dthp polynucleotides of the invention

SQ Sequence 3427 BP; 682 A; 1145 C; 935 G; 665 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1.58e-147 Length: 3427  
 Score: 1372.00 Matches: 265  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.64% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-778-963a-2 (1-266) x ABK71563 (1-3427)  
 QY 1 MetMetLysThrLeuSerSerGlyAsnCySeThrLeuSerValProAlaLysAsnSerTyr 20  
 DB 644 ATGATGAAGACTTGTCTCAGCGGGAATGCACGCTCATGTGCGCCGCAAAATCTCATAC 703  
 QY 21 ArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40  
 DB 704 CGCATGGTGGTGGTGGGCTCTCGGGTGGGCAAGAGCTTCATCGTCTCGCTTCTC 763  
 QY 41 AsnGlyArgPheGluAaPrgIntYrThrProThrIleGluAaPheHISArgValTyr 60  
 DB 764 AATGGCCGCTTGGAGACAGTACACACCACTCAGAGACTTCCACCGTAAGGTATAC 823  
 QY 61 AsnIleArgGlyAaPmetTyrGlnLeuAAspIleLeuAaPThrSerGlyAsnHISProPhe 80  
 DB 824 AACATCCGCGCGACATGATACAGCTGCATCTGATATCTTGCAACACACCTTC 883  
 QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 DB 884 CCGCCATGGCGCAAGCTGTCTCATCTCAGAGGATGCTTCATCGTGTTCAGCTG 943  
 QY 101 AspAsnArgGluSerPheAaPrgValValArgLeuGlnLysGlnIleLeuGlnValLys 120  
 DB 944 GATTAACCGGAGTCTTGCATGAGGTCAAGGCGCTTCAAGACGATCTGGAGGTCAAG 1003  
 QY 121 SerCysLeuLysAsnLysThrLysGluAlaAlaGluLeuPProMetValIleCysGlyAsn 140  
 DB 1004 TCTGCTGAAAGAACAGACCAAGAGGCGGAGGCTGCCATGTCATTTGTGGCAAC 1063  
 QY 141 LysAsnAAspHISGlyLysLeuCyAsrGlnValProThrThrGlnValGluLeuVal 160  
 DB 1064 AAGAACGACCAACGCGAGCTGTGCGCAGGTGCCACCAACGAGCCGAGCTCGTGTG 1123  
 QY 161 SerGlyAaPrgLysAsnSerIleArgPheGluValSerAlaLysLysAsnThrAsnValAaP 180  
 DB 1124 TCGGGCGACGAGAACTGCGCTTCTGAGGTGTGCGCAAGAAACACCAAGGTGAC 1183  
 QY 181 GluMetPheTyrValIleuPheSerMetAlaLysLeuProHISGluMetSerProAlaLeu 200  
 DB 1184 GAGATGTTCTAAGTCTTCAAGCATGACCAAGCTGCACACGAGATGAGCCCGCCCTG 1243  
 QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHISProArgProPheCysMetArg 220  
 DB 1244 CATGCAAGATCTCCGTCGATGAGTGAAGCCCTTCCACCCCAAGCCCTTGTGATGGC 1303  
 QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgProSerVal 240  
 DB 1304 CGCGTCAAGAGATGAGACGCTTATGATGATGTCTCGCCCTTCCGCCGCCCAAGCTTC 1363  
 QY 241 AsnSerAspLeuLysTyrIleValAlaLysValLeuArgGluGlnIleAlaArgGluArg 260  
 DB 1364 AACAGTACCTCAAGTACATCAAGGCAAGGTCTTCGGGAAGGCCAGGCCGTGAGAGG 1423  
 QY 261 AspLysCyThrIleGln 266  
 DB 1424 GACAAAGTGCACCATCCAG 1441  
 RESULT 8  
 ABL92134  
 ID ABL92134 standard; cDNA; 3020 BP.  
 XX  
 AC ABL92134;

XX 30-MAY-2002 (first entry)  
DT Mouse Tumour Endothelial Marker polynucleotide SEQ ID NO 292.  
XX  
DE Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytoskeletal;  
KW normal endothelial marker; pan-endothelial marker; immunostimulant;  
KW antiangiogenic; tumour; neovascularisation; vascularised tumour;  
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;  
KW psoriasis; gene; ss.  
XX  
OS Mus musculus.  
XX  
PN WO200210217-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 01-AUG-2001; 2001WO-US024031.  
XX  
PR 02-AUG-2000; 2000US-0222599P.  
PR 11-AUG-2000; 2000US-0224360P.  
PR 11-APR-2001; 2001US-0282850P.  
XX  
PA (UYJO ) UNIV JOHNS HOPKINS.  
XX  
PI St Croix B, Kinzler KW, Vogelstein B;  
XX  
DR WPI; 2002-291856/33.  
DR P-PSDB; ABB90781.  
XX  
XX An isolated molecule comprising an antibody variable region which  
PT specifically binds to an extracellular domain of a tumor endothelial  
PT marker (TEM) protein, useful for inhibiting tumor growth.  
XX  
XX Disclosure: Page 294-295; 331pp; English.  
XX  
XX The invention relates to an isolated molecule comprising an antibody  
CC variable region which specifically binds to an extracellular domain of a  
CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,  
CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM  
CC proteins have cytostatic, immunostimulant and antiangiogenic activity.  
CC They are useful for inhibiting tumour growth, neovascularisation in subjects  
CC bearing a vascularised tumour, polycystic kidney disease, diabetic  
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM  
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)  
CC are disclosed, as are marker oligonucleotide sequences: tumour  
CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal  
CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers  
CC (PEM) ABL91903-ABL91995  
XX  
SQ Sequence 3020 BP; 690 A; 848 C; 780 G; 702 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 9, 21e-140 Length: 3020  
Score: 1304.00 Matches: 252  
Percent Similarity: 96.24% Conservative: 4  
Best Local Similarity: 94.74% Mismatches: 10  
Query Match: 94.70% Indels: 0  
Gaps: 0

US-09-778-963A-2 (1-266) x ABL92134 (1-3020)

QY 1 MetMetLysThrLeuSerSerGlyValSerValProAlaLysAsnSerTyr 20  
DB 376 ATGATGAAGACCTTGTCCAGTGGAACTGCACACTCAATGCTCTAAGAACTCTAC 435  
QY 21 ArgMetValIleuGlyValAsnArgValGlyLysSerSerIleValSerArgPheLeu 40  
DB 436 CGCATGCTGCTGCTGGGCTCTCCGAGTGGGCAAGAGCTCAATGTTCTCCGCTTCTC 495  
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60  
DB 496 AATGACCTCTTGAAGACCAAGTACAGCCCACTATGAGAGACTTTCATGCAAGGTGAC 555

QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
DB 556 AACATCCACGGGAGCATTTACAGCTGATATCTCTGACACCTCTGGACACCACTTC 615  
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
DB 616 CTGTCATGCGCCGCTCTCCATCTCTACAGAGATGCTTCACTCTGCTGCTGAGCTG 675  
QY 101 AspaAsnArgGluSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluValLys 120  
DB 676 GATAGCCGGAGGTCTTGTGATGAGGTCAAGCGCTCCGAAACAGATCTGAGGTCAAG 735  
QY 121 SerCysLeuLysAsnLysThrLysGluValIleAlaGluLeuProMetValIleCysGlyAsn 140  
DB 736 TCTGCTCGAAGATTAACCAAGAGGACAGACAGCTGCCCATGTGATCTGTGGAGAC 795  
QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluValGluLeuVal 160  
DB 796 AAGATGACCAACAGTGAAGCTGCGCCAGGCTCCCTGATGAGAGCTGAGCTGTG 855  
QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180  
DB 856 TCTGTGATGAACAACTGGCTTATTTGAGGTGTCAAGCAAGAAAGAACTAATGTGAAC 915  
QY 181 GluMetPheTyrValIleuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200  
DB 916 GAGATGTTCTATGCTGTCTTACGATGACCAAGCTGCCCATGATGAGCTGCTGACTG 975  
QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
DB 976 CACCATTAAGATCTCGTCAAGTACAGGAGCTCTTTCACCCCGGCTTCTGATGAGCT 1035  
QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240  
DB 1036 CGCACTAAGTTCGACAGGTGCTATGCAATGCAATGCTTTCACCCCTTGGCCGACAGTGTG 1095  
QY 241 AsnSerAspLeuLysTyrIleLysAlaLysValLeuArgGluGlnAlaArgGluArg 260  
DB 1096 AACAGTACCTCAAGTACATCAAGCCCAAGGTCTTCAAGGAGGAGCCGAGAGAGG 1155  
QY 261 AspLysCysThrIleGln 266  
DB 1156 GACAAAGTGTACATCCAG 1173

RESULT 9  
ABX72059  
ID ABX72059 standard; DNA; 3020 BP.  
XX  
AC ABX72059;  
XX  
DT 12-MAR-2003 (first entry)  
XX  
XX DNA encoding mouse tumour endothelial marker mTEM 2.  
XX  
DE Mouse; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
KW Tumour endothelial marker; normal endothelial marker; PEM;  
KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
KW neovascularisation; immune response; cytoskeletal; antidiabetic; gene;  
KW ophthalmological; antirheumatic; antirheumatic; antipsoriatic; ds.  
XX  
OS Mus sp.  
XX  
PN WO200283874-A2.  
XX  
PD 24-OCT-2002.  
XX  
PF 10-APR-2002; 2002WO-US008253.  
XX  
PR 11-APR-2001; 2001US-0282850P.  
PR 06-FEB-2002; 2002US-0354262P.  
XX

PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Carson-Walter E, St Croix B, Kinzler KM, Vogelstein B;  
 XX  
 DR WPI; 2003-093016/08.  
 XX P-PSDB; AB054488.  
 XX  
 PT New purified human transmembrane protein, designated as tumor endothelial  
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,  
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.  
 XX  
 PS Disclosure; Page 329-331; 374pp; English.  
 XX  
 CC The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumor ECs. Tumor endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumors as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for  
 CC inducing an immune response to tumour endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumours. The present  
 CC sequence represents a mouse TEM gene  
 XX  
 SQ Sequence 3020 BP; 690 A; 848 C; 780 G; 702 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 9,21e-140 Length: 3020  
 Score: 1304.00 Matches: 252  
 Percent Similarity: 96.24% Conservative: 4  
 Best Local Similarity: 94.74% Mismatches: 10  
 Query Match: 94.70% Indels: 0  
 DB: 7 Gaps: 0  
 US-09-778-963A-2 (1-266) x ABX72059 (1-3020)  
 QY 1 MetMetlysthrleuSerSerGlyAsnCysThrleuSerValProAlaLysAsnSerTyr 20  
 Db 376 ATGATGAGAACCTTGTCCAGTGGGAACTGCACACTCATGTGCTTAAACCTCCCTAC 435  
 QY 21 ArgMetValValleuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheIleu 40  
 Db 436 CGCATGGTGTGCTGGGCTCCCGAGTGGCAAGACCTCATGTGTCTCCGCTCTCTC 495  
 QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60  
 Db 496 AATGCGCGCTTTGAGACACATACACGCCCATTCAGACCTTCATCCGACAGGTGAC 555  
 QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 Db 556 AACATCCACCGGGGACATGTACACAGCTGATTCCTGGACACCTTGGCAACACCATTC 615  
 QY 81 ProIleMetArgAspGluSerIleLeuThrGlyAspValPheIleLeuValPheSerIleu 100  
 Db 616 CTGTCCCATGCGCGCTCTTCATCTCCACAGAGATGCTTCATCTCGGTGTCAAGCTG 675  
 QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluValLys 120  
 Db 676 GATAGCGCGAGTCTTTGATGAGGTCAAGCGCTCCGAAACAGATCTCTGAGGTCAAG 735  
 QY 121 SerCysLeuLysAsnLysThrLysGluAlaAlaGluLeuPProMetValIleCysGlyAsn 140  
 Db 736 TCCCTGCGCTGAAGATTAACCAAGAGGACACAGAGCTCCCATGTGATCTGTGGGAC 795  
 QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160  
 Db 796 AAGAAATGACCAAGTGAAGCTGTGCGCAAGGTCTCTGCATGAGGCTGAGCTGTG 855  
 QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180

Db 856 TCTGTGATGATAAACTGCGCTTATTTGAGGTGTCAACCAAGAACTAATATGTAAC 915  
 QY 181 GluMetPheTyrValleuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200  
 Db 916 GAGATGTTCTATGTGCTGTTCAGATGACCAAGCTGTCCATGAGATGAGCCCTGCACTG 975  
 QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
 Db 976 CACATTAAGATCTCCGTCGATGACGGCGATGCTTTTCAACCCCGGCTCTTCATGCGGT 1035  
 QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240  
 Db 1036 CGCATTAAGTTCGCGAGGTGCTTATGAGCATGTCTCACCTTTGGCCCAACGCGCCAGTTC 1095  
 QY 241 AsnSerAspLeuLysTyrIleLysAlaLysValleuArgGluGlnAlaArgGluArg 260  
 Db 1096 AACGTATACCTCAAGTACATCAAGCCCAAGGTCTTACGGAGGACAGCCCGAGAGAG 1155  
 QY 261 AspLysCysThrIleGln 266  
 Db 1156 GACAAAGTATGACATCCAG 1173  
 RESULT 10  
 ID AAS90571 standard; cDNA; 951 BP.  
 XX  
 AC AAS90571;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #26375.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 ER 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSB-) HYSBQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR P-PSDB; ABG26384.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1; SEQ ID NO 26375; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 951 BP; 202 A; 291 C; 293 G; 165 T; 0 U; 0 Other;

### Alignment Scores:

Pred. No.:	6,296-120	Length:	951
Score:	1126.00	Matches:	225
Percent Similarity:	91.87%	Conservative:	4
Best Local Similarity:	90.24%	Mismatches:	20
Query Match:	81.77%	Indels:	0
DB:	5	Gaps:	0

US-09-778-963A-2 (1-266) x AAS90571 (1-951)

QY	21	ArgMetValValIleuGIyAlaSerArgValAlGIyLysSerSerIleValAlaSerArgPheLeu	40
Db	142	GCGATGGTGTCTGGGTCTCTCGAGTGGGCAAGACTTCATCGTGTCTCGCTTCTC	201
QY	41	AsnGIyArgPheGIyAspGIyInTyrrThrProthrlIeGIyAspPheHisArgLysValTy	60
Db	202	AATGGCCGCTTTGAGGACCAAGTACACACCACCATGAGGACTTCCACCGTAAGTTAC	261
QY	61	AsnIleArgGIyAspMetTyrgIleuAspIleLeuAspThrSerGIyAsnHisProPhe	80
Db	262	AACATCCCGGCGGACATGTATCCAGCTCCGATCTGTGATATCTCTGGCAACACCCCTTC	321
QY	81	ProAlaMetArgArgIleuSerIleLeuThnGIyAspValPheIleLeuValPheSerLeu	100
Db	322	CCCGGCATCGGACGAGCTGTTCATCTCCACAGGAGGAGGCCCATGTGTGCTGGGCTGGGGC	381
QY	101	AspAsnArgGIySerPheAspGIyValLysArgIleuGIyInLysGIyInIleLeuGIyValLys	120
Db	382	GGCAGGGCCGACGGGCATGGGTGGGAGTGTCTGTGGGCACTTAGAGAGATCCTCGAGAGTCAAG	441
QY	121	SerCysLeuLysAsnLysThryLysGIyValAlaAlaGIyLeuProMetValIleCysGIyAsn	140
Db	442	TCCTGCTGTGAAGAACAAAGACCAAGAGAGGGCGGAGCTGGCCATGTGCATCTGTGGCAAC	501
QY	141	LysAsnAspHisGIyGIyIleuCysArgGIyValProThrThrGIyValaGIyLeuLeuVal	160
Db	502	AAGAACGACACAGCGGACGCTGTGGCCGACAGTGGCCACACGAGGCGGAGCTGTGGTG	561
QY	161	SerGIyAspGIyAsnSerAlaTyrrPheGIyValSerAlaLysLysAsnThrAsnValAsp	180
Db	562	TCGGGCGACGAGAACCTGGCGCTTACCTTGGAGGTTCGGCCCAAGAAACACCAACGTGGAC	621
QY	181	GIuMetPheTyrrValIleuPheSerMetAlaLysLeuProHisGIuMetSerProAlaLeu	200
Db	622	GAGATGTTCTTACGTGCTCTTACAGCATGGCCAAAGCTGGCACACAGATGAGACCCCGCCTG	681
QY	201	HisAsnArgLysIleSerValGIyInTyrrGIyAspAlaPheHisAspProArgProPheCysMetArg	220
Db	682	CATCCCAAGATTCCTCGGTGAGTACGGGTGACGCTTCCACCCCAAGGCCCTTCTGCATGGCGC	741
QY	221	ArgValLysGIuMetAspAlaTyrrGIyMetValSerProPheAlaArgArgProSerVal	240
Db	742	CGCGTCAGAGATGACCGCTTATGGCAATGTCTTGGCCCTTGGCCGCGCCCAAGCGTC	801
QY	241	AsnSerAspLeuLysTyrrIleLysAlaLysValIleuArgGIyInGIyInAlaArgIyValArg	260
Db	802	AACAGTGAACCTCAAGTATCAAGAGGCCAAGGAGTCTTCCGGAAGGCCAGGCCGCTGAGAGG	861
QY	261	AspLysCysThrlIleGIyIn	266
Db	862	GACAAAGTGCACCATTCAG	879

	RESULT	11
ABSS4740	ID	ABSS4740 standard; DNA; 11221 BP.
XX AC		
XX ABSS4740;		
DT DT	04-DEC-2002	(first entry)
XX DE		
DE Genomic DNA encoding human Ras-like protein.		
XX XX		
KW KW	Human; Ras-like; cancer; apoptosis; cell proliferation; AIDS; allergy/	
KW KW	acquired immunodeficiency syndrome; Alzheimer's disease; osteoporosis;	
KV KV	Parkinson's disease; aplastic anaemia; atherosclerosis; inflammation;	
KX KW	bacterial infection; viral infection; fungal infection; gene therapy;	
KW KW	parasitic infection; protozoal infection; helminthic infection; SNP;	
XX KW	Ras-Inhibitor; Ras-Stimulator; single nucleotide polymorphism; gene; d	
OS OS	Homo sapiens.	
XX XX		
FH FH	Location/Qualifiers	
FT FT	CDS	3000..8223
FT FT	/tag= a	
FT FT	/product= "human Ras-like protein"	
FT FT	/number= b	3000..3270
FT FT	/tag= c	
FT FT	/number= 1	3271..7693
FT FT	/tag= c	
FT FT	/number= 1	
FT FT	replace(3951,T)	
FT FT	/tag= d	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(4127,T)	
FT FT	/tag= e	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(4157,A)	
FT FT	/tag= f	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(4513,T)	
FT FT	/tag= g	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(6594,C)	
FT FT	/tag= h	
FT FT	exon	7694..8220
FT FT	/tag= i	
FT FT	/number= 2	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(8409,C)	
FT FT	/tag= j	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(8437,G)	
FT FT	/tag= k	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(8579,T)	
FT FT	/tag= l	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(10292,G)	
FT FT	/tag= m	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
FT FT	replace(10792,G)	
FT FT	/tag= n	
FT FT	/standard name=	"Single nucleotide polymorphism (SNP) "
XX XX		
NN NN	WO200262849-A2.	
PD PD	15-AUG-2002.	
PP PP	28-JAN-2002; 2002WO-US002301.	
PR PR	08-FEB-2001; 2001US-00778963.	
PA PA	(PEKE ) PE CORP NY.	

PI	Needham B Ketchum KA, Di Francesco V, Beasley EM;
DR	WPI; 2002-706901/76.
XX	P-PSDB; AEG70951.
PT	New human Ras-like polypeptides, useful for treating diseases associated
PT	with an increase in apoptosis or cell proliferation e.g., cancer.
XX	Claim 4; Fig 3; 73pp; English.
XX	The present invention relates to a new polypeptide comprising a fully
CC	defined 266-amino acid sequence or its fragment, orthologue or allelic
CC	variant. The allelic variant or orthologue is encoded by a nucleic acid
CC	molecule that hybridises under stringent conditions to the opposite
CC	strand of a nucleic acid molecule with a fully defined sequence of 3082
CC	or 11221 bp. The pharmaceutical composition comprising the agent that
CC	binds to the polypeptide is useful for treating a disease or condition
CC	mediated by a human Ras-like protein e.g. cancer. Likewise, the
CC	polypeptide is useful for identifying modulators of Ras-like protein
CC	activity or expression by screening a compound for the ability to
CC	stimulate or inhibit interaction between the Ras-like protein and an
CC	interacting molecule or Ras-like activity. The polypeptide is useful for
CC	diagnosing a disease, or predisposition to a disease, or treating
CC	diseases associated with an increase in apoptosis or cell proliferation
CC	e.g. AIDS (acquired immunodeficiency syndrome), Alzheimer's disease,
CC	Parkinson's disease, aplastic anaemia, allergies, athroscleorosis,
CC	osteoporosis or bacterial, viral, fungal, parasitic, protozoal or
CC	helminthic infections, and inflammation. The present nucleic acid
CC	sequence encodes the human Ras-like protein of the invention
XX	
SQ	Sequence 11221 BP; 2512 A; 3108 C; 3028 G; 2573 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	1.2e-93 Length: 11221
Score:	911.00 Matches: 176
Percent Similarity:	98.33% Conservative: 1
Best Local Similarity:	97.78% Mismatches: 3
Query Match:	66.16% Indels: 0
DB:	Gaps: 0
US-09-778-963A-2 (1-266) x ABS54740 (1-11221)	
QY	87 SerlleleuthrGlyAspValPheilleleuValPheSerleuAspAsnArgLysSerPhe 106 :::
Dd	7681 TCCTCCTCGCAAGGGATGTCTTCACTCTGCGTTCAACCTGGATAAACCGGAGTCCTTC 7740
QY	107 AspgluValIysArgLeuGlnIleGlnIleleuGluValIysSerCylleuIysAsnIys 126 
Dd	7741 GATGAGGCACAAGCGCTTCAGAAGAGATCTCGAGGCTCAAGTCTTGCTTAAGAAACAG 7800
QY	127 ThrIysGluAlaIleagluLeuProMetValIleCySGIYasnuIyAsnAspHiIGLyLu 146 
Dd	7801 ACCAAGAGGGGGCGGAAGCTGCCATGTGTCATTGTGGCACAGAACACACACGGCGAG 7866
QY	147 LeuCysArgInuValProThrThrluIleagluLeuValSerGIYAspGuAnsSer 166 
Dd	7861 CTGTGCGCCGAGGTGCCACCACCGAGCGGAGTGTGTGTGCGGCGACAGAGACTGC 7920
QY	167 AlaIyrPheGluValSerAlaIyalyAsnThrArnValAspGluMetPheYrTalIeu 186 
Dd	7921 GCCCACTTCGAGGTGTGCGCCAAAGAAACCAACGAGGAGATGTTTCACTGTCTC 7980
QY	187 PheSerMetAlaIyaleuProHIGLyMetSerProAlaIeuHIeArgIyleSerVal 206 
Dd	7981 TTCAGCATGGCCAGCTGCCACACAGATGAGCCCCCGCTGCATCGCAAGATCTCCGTG 8040
QY	207 GlnTyrgIyaAspAlaPheHisProArgProPheCYsmetArgArgValIyglIueMetAsp 226 
Dd	8041 CAGTACGGTGAACGCTTCACCCGAGCGCTTCTGATGCGCGCGGTCAAGAGATGAC 8100
QY	227 AlaTyrgIymetValSerProPheAlaArgArgProSerValAnSerdAspleuIyTr 246 
Dd	8101 GCCTATGGCATAGTGTGCGCTTCGCGCGCGCGCCACAGCTCAACAGTCAAGTAC 8166

Qy	247	11leyysalalyysVal1leuAryGJUG1yG1alAalArgJ1unrAapLySvYstH11LG1n	266
Db	8161	ATCAGGCCAAGGTCCTTCGGAGAGCCAGGCCCTGTAGAGGACACAGTCACATCCAG	8220
RESULT 12			
ID	AAZ36892		
AC	AAZ36892;		
XX			
DT	13-MAR-2000	(first entry)	
XX			
DE	CDNA encoding an activator of G protein signalling (AGS) protein.		
KM	Activator of G protein signalling; AGS; ras-related G protein;		
KM	GTP hydrolysis; G protein activity; pheromone response pathway;		
KW	cellular signal transduction; G-gamma selectivity;		
XX			
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	1..846	
FT		/*tag= a	
FT		/product= "activator of G protein signalling (AGS)	
FT		protein"	
XX			
PN	W09956670-A1.		
XX			
PD	18-NOV-1999.		
XX			
PE	07-MAY-1999;	99WO-US010151.	
XX			
PR	08-MAY-1998;	98US-0084842P.	
XX			
PR	07-OCT-1998;	98US-0103355P.	
XX			
PA	(CADU-) CADUS PHARM CORP.		
XX			
PI	Ciemowski M, Duzic E;		
XX			
DR	WPI; 2000-072337/06.		
XX			
DR	P-PSDB; AAY53921.		
XX			
PT	A new activator of G protein signalling used to treat disorders		
XX			
PT	characterized by an aberrant AGS protein activity.		
XX			
PS	Claim 3; Fig 3A; 162pp; English.		
XX			
CC	The present sequence encodes an activator of G protein signalling (AGS)		
XX			
CC	protein. The cDNA sequence was isolated from a human liver cDNA library.		
XX			
CC	The AGS protein exhibits homology to ras-related G proteins, and contains		
XX			
CC	alterations in conserved amino acids consistent with a deficiency in GTP		
XX			
CC	hydrolysis activity. AGS stimulates G protein activity, G protein-coupled		
XX			
CC	signal transduction and the pheromone response pathway in a receptor-		
XX			
CC	independent manner. The AGS protein also shows G-gamma selectivity, as		
XX			
CC	measured by growth assays in yeast expressing various mammalian G-gamma		
XX			
CC	constructs, and tissue-specific expression, as measured by Northern blot		
XX			
CC	analysts. The AGS protein can be used to screen for compounds that		
XX			
CC	modulate cellular signal transduction. The protein is used to treat		
XX			
CC	disorders characterized by an aberrant AGS protein activity or AGS		
XX			
CC	nucleic acid expression		
XX			
SQ	Sequence 846 BP; 176 A; 288 C; 252 G; 130 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. NO.:	1.64e-91	Length:	846
Score:	879.00	Matches:	172
Percent Similarity:	77.70%	Conservative:	44
Best Local Similarity:	61.87%	Mismatches:	46
Query Match:	63.83%	Indels:	16
DB:	3	Gaps:	4





```

Db      401 CCCGCCATGCGGCGCTCTCCATCTCAGAGAGCGTTTTCATCTGTTGAGTCTG 460
Qy      101 AspaanargluserpheasgluVallysrghleuglnlysgnileuVallys 120
      461 GACACCGGCGCTCTTGGAGAGGTGCGCGCTCAGCGAGCGAGATCTTGACACCAAG 520
Qy      121 SerCysleuylsasnlysrThlysgluValalaglueupProwetValleCysglYasn 140
      521 TCTTGCTCAAGAACAAACCAAGAGAACCTGACGTCCTCGTATCTTGCGGCAAC 580
Qy      141 LysasnaphisglYleuValCysargGlnValProthrThrgluValgluLeuVal 160
      581 AAGGATGACCGC---GACTTCTACCGCGAGGTGACCGACCGAGATGACGACCTG 637
Qy      161 SerGlyasp---GluasnserAlaTyrPheGluValSerAlaValylsAsnThrAsnVal 179
      638 GCGCAGACACCCCGCGCGCTGCGCTTACGAGATCTCGCGCAAGAACAGACGCTG 697
Qy      180 AspGluMetPheTyrValleuPheSerMetAlaLysLeuProHISgluMetSerProAla 199
      698 GACCAAGATGTTCCGCGCGCTCTTCCATGAGCCAAAGCTGCGCAGAGATGACCCAGAC 757
Qy      200 LeuHisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMet 219
      758 CTGCAACCGCAAGGTCGTGTCAGTACTGACGCGTGCACAGAG---GCGCTG 811
Qy      220 ArgArgVallysgluMet-----AspAla 227
      812 CGGAAACAGAGCTGCTGCGGCGCGCAGCGCGCGCGCGCGAGACCCGCGCGAGCGCC 871
Qy      228 TyrGlyMetValSerProPheAlaArgArgProSerValaenseArgLeuValTyr 247
      872 TTTCGCAATCGTGCAACCTTTCGCGCGCGCGCGCGAGCGTACACGACCTTATGTATC 931
Qy      248 LysAlaLysValleuArgGlnGlyGlnAlaArgGluArgAspLysCysThrIle 265
      932 CGCGAGAAAGCGCGCGCGCGCGCGAGCGCCAAAGAGCAAGAGCGCTGCTGATC 985
Db
RESULT 14
ACC42325
ID ACC42325 standard; CDNA; 1744 BP.
XX
AC ACC42325;
XX
XX 22-MAY-2003 (first entry)
XX
DE Human MAP kinase cascade activator #44 CDNA.
XX
KW Human; Elki phosphorylation; Elki phosphorylation kinase; virucide;
KW antiinflammatory; immunomodulator; cytostatic; antiallergic; anti-HIV;
KW antineumatic; antidiabetic; antiaesthetic; antiaesthetic; gene therapy;
KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
KW Iga nephritis; gene; ss.
XX
OS Homo sapiens.
XX
PN W02003008589-A1.
XX
PD 30-JAN-2003.
XX
PF 15-JUL-2002; 2002W0-JP007174.
XX
PR 18-JUL-2001; 2001JP-00218204.
PR 31-AUG-2001; 2001JP-00263450.
PR 21-JAN-2002; 2002JP-00012176.
XX
PA (ASAH) ASAH KASEI KOGYO KK.
XX
PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
XX
DR WPI: 2003-229582/22.
DR P-PSDB; ABR41059.

```

```

XX      XX      Elki phosphorylation-associated gene and its encoded protein with MAP
PT      PT      kinase cascade effect, applicable in diagnosis of and developing drugs
PT      PT      for e.g. inflammations, autoimmune diseases, viral diseases and cancer.
XX      XX      Claim 4; Page 316-320; 762pp; Japanese.
XX      XX      The invention relates to a novel purified protein having Elki
CC      CC      phosphorylation activity and/or an activity of activating Elki
CC      CC      phosphorylation kinase. A protein of the invention has antiinflammatory,
CC      CC      immunomodulator, virucide, cytostatic, antiallergic, antineumatic,
CC      CC      antidiabetic, antiaesthetic, antiaesthetic, and anti-HIV activity. The
CC      CC      polynucleotides may have a use in gene therapy. The gene and its encoded
CC      CC      protein are applicable in diagnosis of and developing drugs for e.g.
CC      CC      inflammations, autoimmune diseases, viral diseases and cancer such as
CC      CC      rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
CC      CC      hepatitis and Iga nephritis. The present sequence is used in the
XX      XX      exemplification of the invention
SQ      SQ      Sequence 1744 BP; 372 A; 575 C; 503 G; 294 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4,47e-91 Length: 1744
Score: 879.00 Matches: 172
Percent Similarity: 77.70% Conservative: 44
Best Local Similarity: 61.87% Mismatches: 46
Query Match: 63.83% Indels: 16
Gaps: 4
US-09-778-963A-2 (1-266) x ACC42325 (1-1744)
Qy      1 MetMetLysThrIleuSerSerGlyAsnCysThrLeuSerValProAlaLysAsnSerTyr 20
      230 ATGATCAAGAGAGATGTCGCCGAGCGACTCGAGCTGAGTATCCCGCAAGAACTGCTAT 289
Db      21 ArgMetValleuGlnAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40
      290 CGCATGTATCTCTCGGCTGCTGTCAGAGTGGGCAAGCGGCATCGTGGCTTCTC 349
Qy      41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60
      350 ACCGCGCGCTTTCAGAGAGCGCTTACACGCTACCATCGAGATCCACCCCAAGTTCTAC 409
Db      61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
      410 TCCATCGCGCGAGGCTCTACAGCTGACATCTTCAACAGTCTGCGCAACACCGCTTC 469
Qy      81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
      470 CCGCGCATGCGGCGCTCTCTCATCTTCAAGAGAGCTTTTCACTCGGTTCAGTCTG 529
Db      101 AspaanargluserpheasgluVallysrghleuglnlysgnileuVallys 120
      530 GACACCGGCGCTCTTGGAGAGGTGCGCGCTCAGCGAGCGAGATCTTGACACCAAG 589
Qy      121 SerCysleuylsasnlysrThlysgluValalaglueupProwetValleCysglYasn 140
      590 TCTTGCTCAAGAACAAACCAAGAGAACCTGACGTCCTCGTATCTTGCGGCAAC 649
Db      141 LysasnaphisglYleuValCysargGlnValProthrThrgluValgluLeuVal 160
      650 AAGGATGACCGC---GACTTCTACCGCGAGGTGACCGACCGAGATGACGACCTG 706
Qy      161 SerGlyasp---GluasnserAlaTyrPheGluValSerAlaValylsAsnThrAsnVal 179
      707 GCGCAGACACCCCGCGCGCTGCGCTTACGAGATCTCGCGCAAGAACAGACGCTG 766
Db      180 AspGluMetPheTyrValleuPheSerMetAlaLysLeuProHISgluMetSerProAla 199
      767 GACCAAGATGTTCCGCGCGCTTTCGCGCGCGCGCGAGTGCAGCGAGAGAGCGCGAC 826
Qy      200 LeuHisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMet 219

```



Job time : 445 secs

---

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

Title: , US-09-778-963A-1

Sequence: 1 ggcgctcgccgcgcggggagaa...aaaaaaaaaaaaaaaaaa 3082

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Pre-processing:	Minimum Match	0%
Post-processing:	Minimum Match	0%

Listing first 45 summaries

Database

1: genbmbl: \*  
 2: 1: gb ba: \*  
 3: 2: gb hcg: \*  
 4: 3: gb in: \*  
 5: 4: gb om: \*  
 6: 5: gb ov: \*  
 7: 6: gb pat: \*  
 8: 7: gb ph: \*  
 9: 8: gb pl: \*  
 10: 9: gb pr: \*  
 11: 10: gb ro: \*  
 12: 11: gb srg: \*  
 13: 12: gb sy: \*  
 14: 13: gb un: \*  
 15: 14: gb vl: \*  
 16: 15: em ba: \*  
 17: 16: em fun: \*  
 18: 17: em hum: \*  
 19: 18: em in: \*  
 20: 19: em nm: \*  
 21: 20: em om: \*  
 22: 21: em ov: \*  
 23: 22: em ox: \*  
 24: 23: em pat: \*  
 25: 24: em ph: \*  
 26: 25: em pl: \*  
 27: 26: em ro: \*  
 28: 27: em srg: \*  
 29: 28: em un: \*  
 30: 29: em vl: \*  
 31: 30: em hcg hum: \*  
 32: 31: em hcg inv: \*  
 33: 32: em hcg other: \*  
 34: 33: em hcg mus: \*  
 35: 34: em hcg pln: \*  
 36: 35: em hcg rod: \*  
 37: 36: em hcg mam: \*  
 38: 37: em hcg wrt: \*  
 39: 38: em hcg: \*  
 40: 39: em hngo hum: \*  
 41: 40: em hngo mus: \*  
 42: 41: em hngo other: \*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	3082	100.0	3082	6	AX572848	AX572848 Sequence
2	3023.8	98.1	3049	6	BC013419	BC013419 Homo sapi
3	2949.4	97.3	2973	6	AX393267	AX393267 Sequence
4	2937.2	95.3	3427	6	AX721069	AX721069 Sequence
5	2806.4	91.1	2832	6	AX393244	AX393244 Sequence
6	2806.4	91.1	2832	9	AF278143	AF278143 Homo sapi
7	2651.6	86.0	2699	9	HSMB03172	HSMB03172 Homo sapi
8	2500.4	81.1	11221	6	AX572850	AX572850 Sequence
9	2500.4	81.1	114771	9	HS566D19	HS566D19 Human DNA
10	1239.8	42.7	178071	2	AC145536	AC145536 Lemur cat
11	1008.4	32.2	17022	6	AX597851	AX597851 Sequence
12	919.4	29.8	274402	2	AC145758	AC145758 Microcebu
13	811.2	26.3	3469	10	AF134409	AF134409 Rattus no
14	779.8	25.3	2824	10	BC036988	BC036988 Mus muscu
15	693.2	22.5	3020	6	AX393362	AX393362 Sequence
16	475.4	15.4	286789	2	AC133723	AC133723 Rattus no
17	464.8	15.1	211075	10	AC0766974	AC0766974 Mus muscu
18	463.6	15.0	478	6	BD108232	BD108232 EST and e
19	463.6	15.0	478	6	BD108232	BD108232 EST and e
20	416	13.5	1744	6	BD190304	BD190304 Elk1 phos
21	416	13.5	1758	6	BC018041	BC018041 Homo sapi
22	416	13.5	1841	6	AR236058	AR236058 Sequence
23	414.4	13.4	1776	6	BD248751	BD248751 GTPase as
24	409.8	13.3	1740	6	AX780438	AX780438 Sequence
25	409.8	13.3	1740	9	AF069506	AF069506 Homo sapi
26	407.2	13.2	979	9	AF172846	AF172846 Homo sapi
27	405.8	13.2	846	9	AF498923	AF498923 Homo sapi
28	405.8	13.1	1746	9	AF177335	AF177335 Homo sapi
29	402.6	13.1	1187	9	AF153192	AF153192 Homo sapi
30	395.4	12.8	217841	2	AC095995	AC095995 Rattus no
31	385	12.5	1390	5	BC056272	BC056272 Datto rer
32	363.8	11.8	210264	2	AC145933	AC145933 Gallus ga
33	362.6	11.8	2294	10	BC026377	BC026377 Mus muscu
34	346.6	11.2	1616	10	AF239157	AF239157 Rattus no
35	346.6	11.2	1689	6	AR236060	AR236060 Sequence
36	343.4	11.1	1623	10	AF009246	AF009246 Mus muscu
37	341.8	11.1	1632	10	BC034166	BC034166 Mus muscu
38	226.2	7.7	5141	9	AF262016	AF262016 Homo sapi
39	234.6	7.6	3986	6	AR236059	AR236059 Sequence
40	234.6	7.6	4990	9	AF222979	AF222979 Homo sapi
41	234.6	7.6	141048	9	AC073621	AC073621 Homo sapi
42	234.6	7.6	183334	9	AC020558	AC020558 Homo sapi
43	235.6	7.6	183427	2	AC090608	AC090608 Homo sapi
44	219.6	6.5	3079	6	AC109402	AC109402 Rattus no
45	199.6	6.5	3079	6	AR236061	AR236061 Sequence

## ALIGNMENTS

RESULT 1				
AX572848				
LOCUS	AX572848	3082 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO02062849.			
ACCESSION	AX572848			
VERSION	AX572848.1	GI:26004937		
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1			
AUTHORS	Neelam, B., Ketchum, K.A., di Francesco, V. and Beasley, E.M.			
TITLE	Isolated human ras-like proteins, nucleic acid molecules encoding them, and uses thereof			



```

Db      1981  GATACCAACGAGCTCGAGGGGACACGATGAGCACAGGCCCACTTTGTCCTTACG 2040
Qy      2041  AAATTCAGGGTACAGCTCCACTTAACACAGGCTCCCTTACTGTCTGTTCTCTCAAG 2100
Db      2041  AAATTCAGGGTACAGCTCCACTTAACACAGGCTCCCTTACTGTCTGTTCTCTCAAG 2100
Qy      2101  CATTATTAAGCACTTACCTGAGGTGCTGAGTTCACTGTGCTTCAAGGAAACCAAGAGGTC 2160
Db      2101  CATTATTAAGCACTTACCTGAGGTGCTGAGTTCACTGTGCTTCAAGGAAACCAAGAGGTC 2160
Qy      2161  CCAAGTCCTGAGCTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2220
Db      2161  CCAAGTCCTGAGCTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2220
Qy      2221  GAGGCGGGGAGAGAGAGCTGGGAGCCAGAACTGAGGCTGGAGGAGTCCGACAGAAAGC 2280
Db      2221  GAGGCGGGGAGAGAGAGCTGGGAGCCAGAACTGAGGCTGGAGGAGTCCGACAGAAAGC 2280
Qy      2281  TCAGGGGGGGGCTCTCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2340
Db      2281  TCAGGGGGGGGCTCTCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2340
Qy      2341  AGGCAATGCTGAGGCTTGAACCCATTAACCAAGGCTGAGGCTGAGGCTGAGGCT 2400
Db      2341  AGGCAATGCTGAGGCTTGAACCCATTAACCAAGGCTGAGGCTGAGGCTGAGGCT 2400
Qy      2401  TATTTTGTATTTTAACTCTCTTAACCAATTAAGGCTTGAAGGCTTGAAGGCTTGA 2460
Db      2401  TATTTTGTATTTTAACTCTCTTAACCAATTAAGGCTTGAAGGCTTGAAGGCTTGA 2460
Qy      2461  GATCTCTAATGAGGCTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2520
Db      2461  GATCTCTAATGAGGCTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2520
Qy      2521  AGCAAGTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2580
Db      2521  AGCAAGTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2580
Qy      2581  GCTCCCTCTCTTGTATGCTTGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2640
Db      2581  GCTCCCTCTCTTGTATGCTTGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2640
Qy      2641  ACCCTTCCCTTGAAGTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2700
Db      2641  ACCCTTCCCTTGAAGTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2700
Qy      2701  TATATCAAGTACACAGATGATACACAGATGATACACAGATGATACACAGATGATAC 2760
Db      2701  TATATCAAGTACACAGATGATACACAGATGATACACAGATGATACACAGATGATAC 2760
Qy      2761  ATACCTGACCTGACACCCAGCCTTGGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2820
Db      2761  ATACCTGACCTGACACCCAGCCTTGGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2820
Qy      2821  CCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Db      2821  CCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Qy      2881  CCTTGGCTGACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Db      2881  CCTTGGCTGACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Qy      2941  TCCGTGTACTATCAATTAAGTGGGTTTGTTCACAAAAA 3000
Db      2941  TCCGTGTACTATCAATTAAGTGGGTTTGTTCACAAAAA 3000
Qy      3001  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
Db      3001  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
Qy      3061  AAAAAAAAAAAAAAAAAAAAAA 3082
Db      3061  AAAAAAAAAAAAAAAAAAAAAA 3082

```

```

Db      3061  AAAAAAAAAAAAAAAAAAAAAA 3082

RESULT 2
BC013419
DEFINITION
Homo sapiens RASD family, member 2, mRNA (cDNA clone MGC:4834
IMAGE:3604196), complete cds.
ACCESSION
BC013419
VERSION
BC013419.2
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
Organism
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 3049)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Dergs, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshibayashi, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3049)
Strausberg, R.
Direct Submission
Submitted (31-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:15426590.
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
http://www.systemsbio.org
contact: amadansystemsbio.org
Anup Madan, Jessica Fahney, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

REMARK
COMMENT
JOURNML
AUTHORS
TITLE
JOURNML
PUBMED
JOURNML
MEDLINE
JOURNML
REFERENCE
JOURNML
AUTHORS
TITLE
JOURNML
REMARK
COMMENT

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAL Plate: 11 Row: D Column: 18

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22027485.

Location/Qualifiers

1..3049

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:4834 IMAGE:3604196"

/tissue\_type="uterus, endometrium adenocarcinoma"

/clone\_id="NIH MGC 44"

/lab\_host="DH10B-R"



gene  
/note="Vector: potb7"  
1..3049  
/gene="RAS2"  
/note="synonyms: TEM2, MGC:4834, Rhes"  
/db\_xref="LocustID:23551"  
208..1008  
CDS  
/codon\_start=1  
/product="RAS family, member 2"  
/protein\_id="AAH13419.1"  
/db\_xref="GI:15426591"  
/db\_xref="LocustID:23551"  
/translation="MKTLSSGNCITLSPAKNSYRMVVLGASRVKSSIVSRFLNGRF  
EDQPTIEDHFRKRIKGMVYLDITDGNHPPFMRRLSLITGVFLVSLDN  
RESFDEVRKLOKOLIFVSKLKNKKEAAELPMYICGNKHGELCRQVPTTELLY  
SGDENCAFEVSAKKNINVDENFYVLFEMALDPEHMSFALRKISVQGDHPRPFC  
MRVYEMDAYGVNPSFARPPVNSDLKIKAKVLEGGARBRDKCTIO"  
259..780  
/note="RAS: Region: Ras subfamily of RAS small GTPases"  
/db\_xref="CDD:smat00173"  
ORIGIN  
Query Match 98.1%; Score 3023.8; DB 9; Length 3049;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3025; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
26 GCCCGCCAGCCCGCCGCTCCGAGACGCGCAGGAGAGATCCCGCCGAGTACCCGGG 85  
23 GCCCGCCAGCCCGCCGCTCCGAGACGCGCAGGAGAGATCCCGCCGAGTACCCGGG 82  
86 AGCCACACAGACTCTGGAGGCTCGGCGGCTGAGACAGCAGCAGCTCCCGCAGCTCC 145  
83 AGCCACACAGACTCTGGAGGCTCGGCGGCTGAGACAGCAGCAGCTCCCGCAGCTCC 142  
146 CGGCGCTTCAGAGCAGCTCTCTGAGCGCTGCGAAGGCCCGCCGCAATTCAGCCCC 205  
143 CGGCGCTTCAGAGCAGCTCTCTGAGCGCTGCGAAGGCCCGCCGCAATTCAGCCCC 202  
206 GAGCCATGATGAAGACTTTTGTCCAGCGGGAATCTGACCGCTGAGTGTCCCGCCAAACT 265  
203 GAGCCATGATGAAGACTTTTGTCCAGCGGGAATCTGACCGCTGAGTGTCCCGCCAAACT 262  
266 CATACCGCAGTGTGCTGGTGCCTTCGGGTGGGCAAGAGCTCCATCTGTCTCGT 325  
263 CATACCGCAGTGTGCTGGTGCCTTCGGGTGGGCAAGAGCTCCATCTGTCTCGT 322  
326 TCCTCAATGCGCGCTTTGAGACACAGTACACACCAACATCGAGAGCTTCCACCTGAAG 385  
323 TCCTCAATGCGCGCTTTGAGACACAGTACACACCAACATCGAGAGCTTCCACCTGAAG 382  
386 TATACCAATCCGCGCGGACATGTACCAAGCTGACATCTGTGATCTCTGGCAACACC 445  
383 TATACCAATCCGCGCGGACATGTACCAAGCTGACATCTGTGATCTCTGGCAACACC 442  
446 CTTCCCGCGCAGTGGCAGGCTGTCCATCTCAAGGGAGTCTTCACTGTGTTCA 505  
443 CTTCCCGCGCAGTGGCAGGCTGTCCATCTCAAGGGAGTCTTCACTGTGTTCA 502  
506 GCTTGATTAACCGGAGATCTTTCATGAGGTCAAGCGCTTCAAGAGCAGATCTGAGAG 565  
503 GCTTGATTAACCGGAGATCTTTCATGAGGTCAAGCGCTTCAAGAGCAGATCTGAGAG 562  
566 TCAAGTCTCGCTGAAGAACCAAGAACCAAGAGCGCGGAGCTGCCCAATGATCTGTGG 625  
563 TCAAGTCTCGCTGAAGAACCAAGAACCAAGAGCGCGGAGCTGCCCAATGATCTGTGG 622  
626 GCACAAGAACCAAGAACCAAGAACCAAGAGCTGTGCGCCAGGTGCCCAACCAAGCGAGCTGC 685  
623 GCACAAGAACCAAGAACCAAGAACCAAGAGCTGTGCGCCAGGTGCCCAACCAAGCGAGCTGC 682  
686 TGTGTGTGGGCGACGAACTCGGCTTACTTGGAGGTGTGCGCAAGAAACCAACCAAG 745  
683 TGTGTGTGGGCGACGAACTCGGCTTACTTGGAGGTGTGCGCAAGAAACCAACCAAG 742

745 TGACAGAGATGTTCTACGTGCTCTTTCAGATAGGCGCAAGCTGCCACAGATAGACCCCG 805  
743 TGACAGAGATGTTCTACGTGCTCTTTCAGATAGGCGCAAGCTGCCACAGATAGACCCCG 802  
806 CCTGTGATCGCAAGATCTCCGTGAGTACGAGTGAAGCTTTCACACCGAGGCTCTTGTGA 865  
803 CCTGTGATCGCAAGATCTCCGTGAGTACGAGTGAAGCTTTCACACCGAGGCTCTTGTGA 862  
866 TGCCCGCGCTTCAGAGATAGAGCGCTTATGAGATGCTGCGCTTGGCGCGCGCCCA 925  
863 TGCCCGCGCTTCAGAGATAGAGCGCTTATGAGATGCTGCGCTTGGCGCGCGCCCA 922  
926 GCGTCAACAGTGAAGCTCAAGTATCATCAAGGCAAGTCTTTCGGAAGGCGAGGCGCGTG 985  
923 GCGTCAACAGTGAAGCTCAAGTATCATCAAGGCAAGTCTTTCGGAAGGCGAGGCGCGTG 982  
986 AGAGGACAAAGTGCACCATTCAGTGAAGCGAGAGATGCTGAGGCGGAGCTTGGCCAGTGC 1045  
983 AGAGGACAAAGTGCACCATTCAGTGAAGCGAGAGATGCTGAGGCGGAGCTTGGCCAGTGC 1042  
1046 TTCAAGGAGATGCGCCCGCAGATGCGCATCTGCGCATCTCCCAACCGAGGCGCGGAGCA 1105  
1043 TTCAAGGAGATGCGCCCGCAGATGCGCATCTGCGCATCTCCCAACCGAGGCGCGGAGCA 1102  
1106 GTCTTGTTCACAGACTTGAAGCAACAGACTGAGAGGCGCGGAGGCTGCTCCGACAT 1165  
1103 GTCTTGTTCACAGACTTGAAGCAACAGACTGAGAGGCGCGGAGGCTGCTCCGACAT 1162  
1166 TCGTGTGCTTTCACAGACTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1225  
1163 TCGTGTGCTTTCACAGACTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1222  
1226 CTCCTGTGAGGAGAGCAATCTGAGAGCTCAAGAGTTAGGAGAGACTCAAGTTACA 1285  
1223 CTCCTGTGAGGAGAGCAATCTGAGAGCTCAAGAGTTAGGAGAGACTCAAGTTACA 1282  
1286 CTTCTCTCTCTGAGGAGTGAAGAAATGTTGATCCAGAGGAGTGAAGATTCGTCGCTCA 1345  
1283 CTTCTCTCTCTGAGGAGTGAAGAAATGTTGATCCAGAGGAGTGAAGATTCGTCGCTCA 1342  
1346 TATGAGAGCTCTCTGAGAGCAAGCTTCAGAGTGAAGAAAGCACAGAAAGCAGATGAAGAA 1405  
1343 TATGAGAGCTCTCTGAGAGCAAGCTTCAGAGTGAAGAAAGCACAGAAAGCAGATGAAGAA 1402  
1406 GTCTCTCTCTCTGAGAGTGAAGAAATGTTGATCCAGAGGAGTGAAGATTCGTCGCTCA 1465  
1403 GTCTCTCTCTCTGAGAGTGAAGAAATGTTGATCCAGAGGAGTGAAGATTCGTCGCTCA 1462  
1466 TCCAGAGCTCTCTCTGAGAGTGAAGAAATGTTGATCCAGAGGAGTGAAGATTCGTCGCTCA 1525  
1463 TCCAGAGCTCTCTCTGAGAGTGAAGAAATGTTGATCCAGAGGAGTGAAGATTCGTCGCTCA 1522  
1526 GCACACACAAAGTGTGCGCCAGAGTGTGCTGACATTTAGGAGTGAAGTGTGCTG 1585  
1523 GCACACACAAAGTGTGCGCCAGAGTGTGCTGACATTTAGGAGTGAAGTGTGCTG 1582  
1586 AAGGGGAGTGGGCAACCTCTGAGAGCGAGGCGCAACCTTGAAGACCAAGCGCACTGCTG 1645  
1583 AAGGGGAGTGGGCAACCTCTGAGAGCGAGGCGCAACCTTGAAGACCAAGCGCACTGCTG 1642  
1646 ACCGCTTCTCAGAGCTCTCTCTGAGAGTGTGCTGCGCCGCAAGTGTGCTGCTGCTG 1705  
1643 ACCGCTTCTCAGAGCTCTCTCTGAGAGTGTGCTGCGCCGCAAGTGTGCTGCTGCTG 1702  
1706 TTGAGAGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765  
1703 TTGAGAGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1762  
1766 CTAGTGAATCCCTCTGAGAGGCGCTCAGCAGTTCAGATTCAGAGGCTTCAGAGGCGG 1825  
1763 CTAGTGAATCCCTCTGAGAGGCGCTCAGCAGTTCAGATTCAGAGGCTTCAGAGGCGG 1822  
1826 ACGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1885

```
Db 1823 |ACGCTTCTCTCTGCTGCCAAACGGGTTCCGTGSCCTGTTGCACTAGCAATTTGA 1882
Qy 1886 |CTTCGCCATTGAGCTCCACGGTTTACAGACAAATTTGCAAGCGTGGGTGGGACGA 1945
Db 1883 |CTTCGCCATTGAGCTCCACGGTTTACAGACAAATTTGCAAGCGTGGGTGGGACGA 1942
Qy 1946 |GGACTGCTTTTAAATGCTCCCAATTTACAGAGGATACACGAGCTGGAGGGGAC 2005
Db 1943 |GGACTGCTTTTAAATGCTCCCAATTTACAGAGGATACACGAGCTGGAGGGGAC 2002
Qy 2006 |ACGATGAGACACAGAGCCCACTTTGTCCTTAGCAAAATTCAGGGTACAGCTCCACTAG 2065
Db 2003 |ACGATGAGACACAGAGCCCACTTTGTCCTTAGCAAAATTCAGGGTACAGCTCCACTAG 2062
Qy 2066 |AACGAGGCTGCTCTACTGTCCTTCAAGCATTTTAAAGCACTTCTGAGTGC 2125
Db 2063 |AACGAGGCTGCTCTACTGTCCTTCAAGCATTTTAAAGCACTTCTGAGTGC 2122
Qy 2126 |TGGGTTCACTGTGCTTAAAGAAACCAAGAGGTCCCGCTGCTGCTGCTGCTGCTGCT 2185
Db 2123 |TGGGTTCACTGTGCTTAAAGAAACCAAGAGGTCCCGCTGCTGCTGCTGCTGCTGCT 2182
Qy 2186 |GCTGCCCAACCACTTCTGACACAGAGCGTGGGGAGGCGGGAGAGAGCACTGGAGCC 2245
Db 2183 |GCTGCCCAACCACTTCTGACACAGAGCGTGGGGAGGCGGGAGAGAGCACTGGAGCC 2242
Qy 2246 |CAGAACTGAGCTGGAGAGGATCCGACGAAAGCTCAGGGCGGGCTTCTCTTGTGCC 2305
Db 2243 |CAGAACTGAGCTGGAGAGGATCCGACGAAAGCTCAGGGCGGGCTTCTCTTGTGCC 2302
Qy 2306 |CGGATTTGGGCTAGCTGGGTACCAACATGTAATCTCAGGCAATGGTGGTTTGAACCTA 2365
Db 2303 |CGGATTTGGGCTAGCTGGGTACCAACATGTAATCTCAGGCAATGGTGGTTTGAACCTA 2362
Qy 2366 |AACCAAGGCGCTTGTATCAGCTCTTAAACAATATTTTATTTTATCTCTTAA 2425
Db 2363 |AACCAAGGCGCTTGTATCAGCTCTTAAACAATATTTTATTTTATCTCTTAA 2422
Qy 2426 |CATATTAAGTTTAAAGGCGCTTAAAGAACTTAAAGTCTTCTTCTGAGG 2485
Db 2423 |CATATTAAGTTTAAAGGCGCTTAAAGAACTTAAAGTCTTCTTCTGAGG 2482
Qy 2486 |TTTGAAGAGGTAGTAATCTTCTCAGGTACACAGAAATCTGTGGTGGGAGAAACA 2545
Db 2483 |TTTGAAGAGGTAGTAATCTTCTCAGGTACACAGAAATCTGTGGTGGGAGAAACA 2542
Qy 2546 |AGCTAGGCTGGGATTCAGTAACACAGATGTCCTCTCTTGAATGCTTGGCC 2605
Db 2543 |AGCTAGGCTGGGATTCAGTAACACAGATGTCCTCTCTTGAATGCTTGGCC 2602
Qy 2606 |TGGGCGCTTCAAGGCTTGGGACATCTTGTCTCAACCCCTCTCCTAGATCAGTCTGA 2665
Db 2603 |TGGGCGCTTCAAGGCTTGGGACATCTTGTCTCAACCCCTCTCCTAGATCAGTCTGA 2662
Qy 2666 |GGGTCTCTGATATTTGTGTAACACATGCCATGTATATCAAGTACACAGATGTA 2725
Db 2663 |GGGTCTCTGATATTTGTGTAACACATGCCATGTATATCAAGTACACAGATGTA 2722
Qy 2726 |ACAAGATGTAACATGCTCCAGGCGGAGCTCTGATACCTGACCTGACCCGACCTT 2785
Db 2723 |ACAAGATGTAACATGCTCCAGGCGGAGCTCTGATACCTGACCTGACCCGACCTT 2782
Qy 2786 |GGCCCTGCTGCTGTGTGCTCAAGAGAGAGCTCAACCTGTCTGTCTCTCTCC 2845
Db 2783 |GGCCCTGCTGCTGTGTGCTCAAGAGAGAGCTCAACCTGTCTGTCTCTCTCC 2842
Qy 2846 |CAACCACTGCTGAGCTTCTGAGCAACAGATACCTTGGCTGACCGGTGTGTGCC 2905
Db 2843 |CAACCACTGCTGAGCTTCTGAGCAACAGATACCTTGGCTGACCGGTGTGTGCC 2902
Qy 2906 |GCTCTACCCAGAGACAGCCCGCACCATGATCTCGTATCACTATATAAAGTG 2965
```

```
Db 2903 |GCTCTACCCAGGACAGCCCGCACCATGATCTCGTATCACTATCAATAAAGTG 2962
Qy 2966 |GGTTGTTTACAAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3025
Db 2963 |GGTTGTTTACAAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3022
Qy 3026 |AAAAAAAAAAAAAAAAAAAAAAAAAAAA 3052
Db 3023 |AAAAAAAAAAAAAAAAAAAAAAAAAAAA 3049

RESULT 3
AX393267
LOCUS AX393267 2973 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 197 from Patent WO0210217.
ACCESSION AX393267
VERSION AX393267.1 GI:19701305
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1
AUTHORS St Croix, B., Kinzler, K.W. and Vogelstein, B.
TITLE Endothelial cell expression patterns
JOURNAL Patent: WO 0210217-A 197 07-FEB-2002;
The Johns Hopkins University (US)
FEATURES
source
1..2973
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 95.7%; Score 2949.4; DB 6; Length 2973;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

26 GCGCGCCGACCGCGGCTCCGAGCAGCGGAGGAGATCCCGGCACTGACCCGGG 85
23 GCCCGCGCCAGCCCGGCGTCCGAGCAGCGAGGAGGAGATCCCGGCACTGACCCGGG 82
86 AGCCACCAAGACTCTGAGGAGGCTCGGCGGCTGGAGAGAGAGAGCTCCCGGAGCTCC 145
83 AGCCACCAAGACTCTGAGGAGGCTCGGCGGCTGGAGAGAGAGAGCTCCCGGAGCTCC 142
146 CGGCGCTTCAAGGAGCTCTTGAAGCGGTGCAAGAGCCCGGCCATTTCCAGGCC 205
143 CGGCGCTTCAAGGAGCTCTTGAAGCGGTGCAAGAGCCCGGCCATTTCCAGGCC 202
206 GAGCCATGATGAGACTTTTGTCCAGCGGAACTGCAAGCTCAAGTGTGCGCCAAAACT 265
203 GAGCCATGATGAGACTTTTGTCCAGCGGAACTGCAAGCTCAAGTGTGCGCCAAAACT 262
266 CATACCGCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 325
263 CATACCGCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 322
326 TCTCTCAATGCGCGCTTGAAGGACAGTACACACCATGATGAGAGCTTCCACGTAAG 385
323 TCTCTCAATGCGCGCTTGAAGGACAGTACACACCATGATGAGAGCTTCCACGTAAG 382
386 TATACCAATTCGCGGCGAGATGTAACAGCTCGAATCTTGGATACCTTGGCAACAC 445
383 TATACCAATTCGCGGCGAGATGTAACAGCTCGAATCTTGGATACCTTGGCAACAC 442
446 CTTTCCCGGATGCGGAGCTGTCATTCCTCAAGGGGATGCTTCAATCTGAGTTCA 505
443 CTTTCCCGGATGCGGAGCTGTCATTCCTCAAGGGGATGCTTCAATCTGAGTTCA 502
506 GCTTGAATAACCGGAGTCTTCTGATGAGTCAAGCGCTTTCAGAACAGATCTTGAAG 565
503 GCTTGAATAACCGGAGTCTTCTGATGAGTCAAGCGCGCTTTCAGAACAGATCTTGAAG 562
```

QY	566	TCAAGTCTGCTTGAAGAACAGACCAAGAGAGCGCGCGAGCTGCCCAATGTCATCTG	625
Db	563	TCAAGTCTGCTCTAAGAACCAAGACCAAGAGAGCGCGAGCTGCCCAATGTCATCTG	622
QY	626	GCAACAAGAAAGACCAAGCGAGAGTGTGCGCGCAGAGTGCACCAACGAGCGAGCTTC	685
Db	623	GCAACAAGAAAGACCAACGCGAGAGTGTGCGCGCAGAGTGCACCAACGAGCGAGCTTC	682
QY	686	TGTGTTCGGGCGACGAGAACTCCGCTTACTTCAGAGTGTGTGGCCAAAGAAACCAAG	745
Db	683	TGTGTTCGGGCGAGAGAACTGCCCTTACTTCAGAGTGTGTGGCCAAAGAAACCAAG	742
QY	746	TGAGAGAAATTTCTAAGTGTCTTTCAGAAATGCGCAAGCTGSCACAGAGATGAGCCCG	805
Db	743	TGAGAGAAATTTCTAAGTGTCTTTCAGAAATGCGCAAGCTGSCACAGAGATGAGCCCG	802
QY	806	CCCTGCATCGCAAGATCTCCGTGCAAGTACGGTGAACGCTTTCACCCAGGCGCTTTCGA	865
Db	803	CCCTGCATCGCAAGATCTCCGTGCAAGTACGGTGAACGCTTTCACCCAGGCGCTTTCGA	862
QY	866	TGCGCCGCTCAAGAGATGAGACGCTTATGGCATATGTCTGGCCCTTGGCCGCGGCCCA	925
Db	863	TGCGCCGCTCAAGAGATGAGACGCTTATGGCATATGTCTGGCCCTTGGCCGCGGCCCA	922
QY	926	GCGTCACAGTGAACCTCAAGTACATCAAGGCGCAAGTCTTTCGGAAAGGCGCAGGCGT	985
Db	923	GCGTCACAGTGAACCTCAAGTACATCAAGGCGCAAGTCTTTCGGAAAGGCGCAGGCGT	982
QY	986	AGAGGACAAATGTCACCAATCCAGTAGAGCGAGAGATGCTGTGGGCGGGGCTTGGCCATG	1045
Db	983	AGAGGACAAATGTCACCAATCCAGTAGAGCGAGAGATGCTGTGGGCGGGGCTTGGCCATG	1042
QY	1046	TTCAAGGAGGTGGCCCGCAAGTACCACATGTGCGCATCTCCCAACGAAGCGCCGACGA	1105
Db	1043	TTCAAGGAGGTGGCCCGCAAGTACCACATGTGCGCATCTCCCAACGAAGCGCCGACGA	1102
QY	1106	GTCCTGTTCAAGACCTTAAGGACCCAGACTGAGGCGCCCGGCGCTGGCTCCGACAT	1165
Db	1103	GTCCTGTTCAAGACCTTAAGGACCCAGACTGAGGCGCCCGGCGCTGGCTCCGACAT	1162
QY	1166	TGCTCTGCTTTCACAGTTTCCTGAGATCCGCTGTTCACAGCTCCCTGGTGGTTTCT	1225
Db	1163	TGCTCTGCTTTCACAGTTTCCTGAGATCCGCTGTTCACAGCTCCCTGGTGGTTTCT	1222
QY	1226	CTCTCTGTGGAGGACATCTCTGCAAGCTCAAGATTAGGACAGAGACTCAAGTTACA	1285
Db	1223	CTCTCTGTGGAGGACATCTCTGCAAGCTCAAGATTAGGACAGAGACTCAAGTTACA	1282
QY	1286	CCTTCTCTCTCTGGGATTGGAAGAAATGTATGTCACAGGAGTGAAGATTCGCGCTCA	1345
Db	1283	CCTTCTCTCTCTGGGATTGGAAGAAATGTATGTCACAGGAGTGAAGATTCGCGCTCA	1342
QY	1346	TATGAGAGCTCCGGGACAAGCTCAGATGTAAAGAGACACAGAAAGCGACATGAGAAAG	1405
Db	1343	TATGAGAGCTCCGGGACAAGCTCAGATGTAAAGAGACACAGAAAGCGACATGAGAAAG	1402
QY	1406	GTCCTCTCTCTCTGGCATTAACCCAGCTTGCTTGGGTGGCACTGGAGAACTTCTC	1465
Db	1403	GTCCTCTCTCTCTGGCATTAACCCAGCTTGCTTGGGTGGCACTGGAGAACTTCTC	1462
QY	1466	TCCCAAGCCTTGAACTTTAAGCTCTGATTAGAGTGCCTCTGACACCCCTCCACCCCA	1525
Db	1463	TCCCAAGCCTTGAACTTTAAGCTCTGATTAGAGTGCCTCTGACACCCCTCCACCCCA	1522
QY	1526	GCAACACACAAGTTGGCCCCAGCTTGCGCTGACATTTAGGCGAGTGAATCTGTGTGTG	1585
Db	1523	GCAACACACAAGTTGGCCCCAGCTTGCGCTGACATTTAGGCGAGTGAATCTGTGTGTG	1582
QY	1586	AAGGGGCGTGGCAACCTCTTGAACACAGCCCAACATTAGACACAGCCCAACTCTGTG	1645
Db	1583	AAGGGGCGTGGCAACCTCTTGAACACAGCCCAACATTAGACACAGCCCAACTCTGTG	1642

Qy	1646	ACCGGTTTCTCAAGCTCTCTCTCTAGTCTCTTCGCGCCGACAGTTGTCTTTGTGNG	1705
Db	1643	ACCGGTTTCTCAAGCTCTCTCTCTAGTCTCTTCGCGCGCGACAGTTGTCTTTGTGNG	1702
Qy	1706	TTGACGCTTTTGTGTCACTAGTATAGTATAGTAAATGAAATCATTTGATCTGTAAAGC	1765
Db	1703	TTGACGCTTTTGTGTCACTAGTATAGTATAGTAAATGAAATCATTTGATCTGTAAAGC	1762
Qy	1766	CTAGTACTCTCTCTTGGCCAGGCTCTCAACCAATTCAAGATCCAGGCTCTCACTCCGGG	1825
Db	1763	CTAGTACTCTCTCTTGGCCAGGCTCTCAACCAATTCAAGATCCAGGCTCTCACTCCGGG	1822
Qy	1826	AACGCTTCCTCTCTGCTCCCAAAACAGGTTTCGNGGCTGTTTGCAGCTACATTTGA	1885
Db	1823	ACGCTTCCTCTCTGCTCCCAAAACAGGTTTCGNGGCTGTTTGCAGCTACATTTGA	1882
Qy	1886	CCTCGGCATTGAGCTCCACGGTTTACAGACAATTGCAACAGCGTGGGCTGGCCAGGCCA	1945
Db	1883	CCTCGGCATTGAGCTCCACGGTTTACAGACAATTGCAACAGCGTGGGCTGGCCAGGCCA	1942
Qy	1946	GGACTGCTTTTTTTTAAATGCTTCCATTTCACAGAGATACACCGAGACTCGAGGGGAC	2005
Db	1943	GGACTGCTTTTTTTTAAATGCTTCCATTTCACAGAGATACACCGAGACTCGAGGGGAC	2002
Qy	2006	ACGATGACACACAGGCCCCACCTTTGTCTCCCTACCAAAATTCAGGGTACAGCTCCACTAG	2065
Db	2003	ACGATGACACACAGGCCCCACCTTTGTCTCCCTACCAAAATTCAGGGTACAGCTCCACTAG	2062
Qy	2066	AACGAGCTGCCCTCTACTGTGTCTCTCTCAAGCATTTATTAAGACCTTACCTGAGGTGC	2125
Db	2063	AACGAGCTGCCCTCTACTGTGTCTCTCTCAAGCATTTATTAAGACCTTACCTGAGGTGC	2122
Qy	2126	TGGGTTCACTGTGCTCTAGAAAACAAGAGGGTCCCCAGTCTGGCTCTGCGCCGCTT	2185
Db	2123	TGGGTTCACTGTGCTCTAGAAAACAAGAGGGTCCCCAGTCTGGCTCTGCGCCGCTT	2182
Qy	2186	GCTGCCCAACAACCTTTTGCACACACACAGCGGTGGGAGGCGGGAGAGACGCTGGGAC	2245
Db	2183	GCTGCCCAACAACCTTTTGCACACACACAGCGGTGGGAGGCGGGAGAGACGCTGGGAC	2242
Qy	2246	CAGAACTGAGCGCTGGAGAGGAAATCGACAGAAAAGTCAAGGCGGGTCTTCTCTGTGGC	2305
Db	2243	CAGAACTGAGCGCTGGAGAGGAAATCGACAGAAAAGTCAAGGCGGGTCTTCTCTGTGGC	2302
Qy	2306	CGGGAATGGGCTATCTGAGTACCAACATGTACTCTGAGGCAATGGTGGTTTGAACCATTA	2365
Db	2303	CGGGAATGGGCTATCTGAGTACCAACATGTACTCTGAGGCAATGGTGGTTTGAACCATTA	2362
Qy	2366	AACCAAAGCCCTTGCATCACTGCTTAAACAATATTTTGTATTTTAACTCTCTAAA	2425
Db	2363	AACCAAAGCCCTTGCATCACTGCTTAAACAATATTTTGTATTTTAACTCTCTAAA	2422
Qy	2426	CATATTGAAGTTTAAAGGCCCTTAAGAAACCTTAGATCTTCTAATGGGCTTTCTGAGG	2485
Db	2423	CATATTGAAGTTTAAAGGCCCTTAAGAAACCTTAGATCTTCTAATGGGCTTTCTGAGG	2482
Qy	2486	TTTCAAGAGAGGTTAATTAATCTTCCAGAGTACACAGAAATGTGTGGGTGGACAGAGCA	2545
Db	2483	TTTCAAGAGAGGTTAATTAATCTTCCAGAGTACACAGAAATGTGTGGGTGGACAGAGCA	2542
Qy	2546	AGCTTACGCTGGGCAATTCAGTACATACACATGTGCTCCCTCTTGTATGCTTGGCCCC	2605
Db	2543	AGCTTACGCTGGGCAATTCAGTACATACACATGTGCTCCCTCTTGTATGCTTGGCCCC	2602
Qy	2606	TGGGGCTTTCAGGGCTTTGGGCAATCTTGTCTCAACCTCTCTCCCTAATCAGTCTGTGA	2665
Db	2603	TGGGGCTTTCAGGGCTTTGGGCAATCTTGTCTCAACCTCTCTCCCTAATCAGTCTGTGA	2662
Qy	2666	GGGTTCCTGTAGATTTTGTGTACACATGCTCCATATATACAGTACACACAGATGTAC	2725
Db	2663	GGGTTCCTGTAGATTTTGTGTACACATGCTCCATATATACAGTACACACAGATGTAC	2722
Qy	2726	ACACAGATGTACATGCTCAAGCCCCACAGCTCTGCAATCTGACCTGACCCACAGCTT	2785

Db 2723 ACAGAGATGACATGCTCCAGCCCGAGCTGTGATCTCGACCTGCAACCCAGCCTT 2782  
Qy 2786 GGGCCCTGCTGCTGCTGTGCTCTCAAGAGACAGCTCCAACTGCTGTGCTCCCTTCC 2845  
Db 2783 GGGCCCTGCTGCTGCTGTGCTCTCAAGAGACAGCTCCAACTGCTGTGCTCCCTTCC 2842  
Qy 2846 CACCACTGCTGAGCTTCTGAGACAGACAGTACCTTGGCTGACCGGCTGCTGCTCC 2905  
Db 2843 CACCACTGCTGAGCTTCTGAGACAGACAGTACCTTGGCTGACCGGCTGCTGCTCC 2902  
Qy 2906 GCTCTCAACCCAGGACAGCCCGCCAGCAGTATCTCCGTGTACACTATCAATAAAGT 2965  
Db 2903 GCTCTCAACCCAGGACAGCCCGCCAGCAGTATCTCCGTGTACACTATCAATAAAGT 2962  
Qy 2966 GGTGTTGTTACA 2976  
Db 2963 GGTGTTGTTACA 2973

RESULT 4  
AX721069 3427 bp mRNA linear PAT 07-MAY-2003  
LOCUS AX721069  
DEFINITION Sequence 29 from Patent WO0220754.  
ACCESSION AX721069  
VERSION AX721069.1 GI:30421905  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Lincoln, S.E., Altus, C.M., Dufour, G.E., Chalup, M.S., Hillman, J.L.,  
Jones, A., Yu, J.Y., Wright, R.J., Gietzen, D., Liu, Toomy, F., Yap, P.,  
Dahl, C.R., Momiya, M., Bradley, D., Rohatgi, S., Harris, B.,  
Roseberry, A.M., Gerstin, B.H., Peralta, C.H., David, M., Panzer, S.,  
Flores, V., Daffo, A., Marwaha, R., Chen, A., Chang, S.C. and Imman, R.R.  
TITLE Molecules for diagnostics and therapeutics  
JOURNAL Patent: WO 0220754-A 29 14-MAR-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
source 1..3427  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: LG:337358.1:2000SEP08"

ORIGIN  
Query Match 95.3%; Score 2937.2; DB 6; Length 3427;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2961; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 26 GCCGCGCCAGCCCGCGCTCCGAGCAGCGAGGAGATCCCGCGCAGTGAACCCGG 85  
Db 459 GCGCGCGCCAGCCCGCGCTCCGAGCAGCGAGGAGATCCCGCGCAGTGAACCCGG 518  
Qy 86 AGCCACCAAGAGCTCTGGAGAGCTCGCGCTGAGCAGCAGCAGCTCCCGCAGCTCC 145  
Db 519 AGCCACCAAGAGCTCTGGAGAGCTCGCGCTGAGCAGCAGCAGCTCCCGCAGCTCC 578  
Qy 146 CGGCGCTTCCAGGAGCTCTGAGCGGTGCGAGAGCGCCGCGCATTCAGAGCCCGC 205  
Db 579 CGGCGCTTCCAGGAGCTCTGAGCGGTGCGAGAGCGCCGCGCATTCAGAGCCCGC 638  
Qy 206 GAGCCATGATGAAGACTTGTCTCAGCGGAACTGACGCTCAGTGTGCGCCCAAAACT 265  
Db 639 GAGCCATGATGAAGACTTGTCTCAGCGGAACTGACGCTCAGTGTGCGCCCAAAACT 698  
Qy 266 CATACCGCATGTGTGTGCTGCTGCTCTCGGCTGAGGAGAGCTCATGCTGTGCTCT 325  
Db 699 CATACCGCATGTGTGTGCTGCTGCTCTCGGCTGAGGAGAGCTCATGCTGTGCTCT 758  
Qy 326 TCCTCAATGGCGGCTTTGAGAGCAGTACACACCAACATGAGAGACTTCCACCGTAAG 385

Db 759 TCCTCAATGGCGGCTTTGAGAGCAGTACACACCAACATGAGAGACTTCCACCGTAAG 818  
Qy 386 TATCAACATTCGGCGGAGCATGTGACAGCTCGACATCTGAGATCTTGGCAACACC 445  
Db 819 TATCAACATTCGGCGGAGCATGTGACAGCTCGACATCTGAGATCTTGGCAACACC 878  
Qy 446 CTTTCCCGGATGCGAGGCTGTCCATCTCAGAGGAGATGTCTCATCTGAGTTTCA 505  
Db 879 CTTTCCCGGATGCGAGGCTGTCCATCTCAGAGGAGATGTCTCATCTGAGTTTCA 938  
Qy 506 GCTGTGATTAACCGGAGTCTCTGATGAGTCAAGCCCTTCAAGAGATCTCTGAGG 565  
Db 939 GCTGTGATTAACCGGAGTCTCTGATGAGTCAAGCCCTTCAAGAGATCTCTGAGG 998  
Qy 566 TCAAGTCTGCTGAGAGACAGACAGAGAGCGGCGAGCTGCCATGCTCATCTGTG 625  
Db 999 TCAAGTCTGCTGAGAGACAGACAGAGAGCGGCGAGCTGCCATGCTCATCTGTG 1058  
Qy 626 GCAACAAAGAACAGACAGCGGAGCTGTGCGCCAGAGTCCCAACAGAGCCGAGCTGC 685  
Db 1059 GCAACAAAGAACAGACAGCGGAGCTGTGCGCCAGAGTCCCAACAGAGCCGAGCTGC 1118  
Qy 686 TGTGTGCGGCGAGAGAACTCCGCTACTTTCAGAGTGTGCGCCAGAGAAACACCAAG 745  
Db 1119 TGTGTGCGGCGAGAGAACTCCGCTACTTTCAGAGTGTGCGCCAGAGAAACACCAAG 1178  
Qy 746 TGAACAGATGTTCTAGCTGTCTTTCAGATGAGCCAGAGTCCACACAGATGAGCCCG 805  
Db 1179 TGAACAGATGTTCTAGCTGTCTTTCAGATGAGCCAGAGTCCACACAGATGAGCCCG 1238  
Qy 806 CCTGTGATCGCAAGATCTCCGTGAGTACAGCTTCCACCCAGAGCCCTTTCGCA 865  
Db 1239 CCTGTGATCGCAAGATCTCCGTGAGTACAGCTTCCACCCAGAGCCCTTTCGCA 1298  
Qy 866 TGCCCGCGCTCAAGAGATGAGACCGCTATGAGATGTCTGCGCCCTGCGCCCGCCCA 925  
Db 1299 TGCCCGCGCTCAAGAGATGAGACCGCTATGAGATGTCTGCGCCCTGCGCCCGCCCA 1358  
Qy 926 GCGTCAACAGTACCTCAAGTACATCAAGAGCCAGAGTCTTGGGAAAGCCAGGCCGTG 985  
Db 1359 GCGTCAACAGTACCTCAAGTACATCAAGAGCCAGAGTCTTGGGAAAGCCAGGCCGTG 1418  
Qy 986 AGAGGAGCAAGTGAACCATCAAGTGAAGAGAGTGTCTGGGCGGGCTTGGCCAGTGC 1045  
Db 1419 AGAGGAGCAAGTGAACCATCAAGTGAAGAGAGTGTCTGGGCGGGCTTGGCCAGTGC 1478  
Qy 1046 TTCAGGAGGAGTGGCCAGATGCCCCATGTCGATCTCCCAACGAGGCCCGGAGCA 1105  
Db 1479 TTCAGGAGGAGTGGCCAGATGCCCCATGTCGATCTCCCAACGAGGCCCGGAGCA 1538  
Qy 1106 GTCTTGTTCACAGACTTATGAGCACAGACTGAGAGCCCGCGGCGCTGCGCATCAT 1165  
Db 1539 GTCTTGTTCACAGACTTATGAGCACAGACTGAGAGCCCGCGGCGCTGCGCATCAT 1598  
Qy 1166 TCGTCTGCTTCTCAAGCTTCTCTGAGTCCGCTTGTTCACAGCTCTTGTGTGTTTCT 1225  
Db 1599 TCGTCTGCTTCTCAAGCTTCTCTGAGTCCGCTTGTTCACAGCTCTTGTGTGTTTCT 1658  
Qy 1226 CTCCTGTGAGAGAGACATCTTGAAGCTTCAAGAGTTTGAAGAGACTCAAGTTTACA 1285  
Db 1659 CTCCTGTGAGAGAGACATCTTGAAGCTTCAAGAGTTTGAAGAGACTCAAGTTTACA 1718  
Qy 1286 CTTTCTCTCTCTGAGGTTTGAAGAAATGTTGATCCAGAGGAGTGAAGATTGCTGCTCA 1345  
Db 1719 CTTTCTCTCTCTGAGGTTTGAAGAAATGTTGATCCAGAGGAGTGAAGATTGCTGCTCA 1778  
Qy 1346 TATGAGCTCTCTGAGGACAGCTCAGAGTGAAGAGACAGAGGCTCAGATGAAG 1405  
Db 1779 TATGAGCTCTCTGAGGACAGCTCAGAGTGAAGAGACAGAGGCTCAGATGAAG 1838  
Qy 1406 GTCTCTCTCTCTGAGGATTAACCCAGCTTGTGAGTGTGAGTGTGAGAACTTCTC 1465

Db	1839	GTCTCTCTCTCTCTGGCAATTAACCCAGACCTTGTTTGGGTGGCAGCTGGGAGAACTCTC	1898
Qy	1466	TCCGAGCCCTGCAACTTTACGCTCTGGTTCACTGCTCTTCGACCCCTCTCCACCCCA	1525
Db	1899	TCCGAGCCCTCAACTTTACGCTCTGGTTCACTGCTCTTCGACCCCTCTCCACCCCA	1958
Qy	1526	GCACACACAAAGTTGGCCCCCAGCTGGCTGACATTGAGCCAGTGGACCTGTGTCTG	1585
Db	1959	GCACACACAAAGTTGGCCCCCAGCTGGCTGACATTGAGCCAGTGGACCTGTGTCTG	2018
Qy	1586	AAGGGGGCGTGGCAACACTCTCTAGACCAAGCCACCACTTGAACACGCCCACTTCTG	1645
Db	2019	AAGGGGGCGTGGCAACACTCTCTAGACCAAGCCACCACTTGAACACGCCCACTTCTG	2078
Qy	1646	ACCGGTTCTTCAGCTCTCTCTCTAGGTCCCTCGGCCACACTGTGGCTTGTGTGG	1705
Db	2079	ACCGGTTCTTCAGCTCTCTCTCTAGGTCCCTCGGCCACACTGTGTGTGTGTGG	2138
Qy	1706	TTGCAGCTGTTTCGTGCATGATAGTATAGATAGAAATGAAATCATTTGACTGTAAAGC	1765
Db	2139	TTGCAGCTGTTTCGTGCATGATAGTATAGTATAGAAATGAAATCATTTGACTGTAAAGC	2198
Qy	1766	CTAGTACTCCCTCCTTGGCCAGGCTCTCACCCAGTTCAATTCACGGGCTTCCACCCGGG	1825
Db	2199	CTAGTACTCCCTCCTTGGCCAGGCTCTCACCCAGTTCAATTCACGGGCTTCCACCCGGG	2258
Qy	1826	ACGCTTCTCTCTCTGCTCCCAACAGGGTTCCGTGGCTCGTTTGGACGTAGACATTTGA	1885
Db	2259	ACGCTTCTCTCTCTGCTCCCAACAGGGTTCCGTGGCTCGTTTGGACGTAGACATTTGA	2318
Qy	1886	CCTCCGCAATTAGAGCTCCACGGTTTACACACAAATTTGACAAAGCTGGGGTGGAGGCCA	1945
Db	2319	CCTCCGCAATTAGAGCTCCACGGTTTACACACAAATTTGACAAAGCTGGGGTGGAGGCCA	2378
Qy	1946	GGACGTCTTTTTTTTAATGCTCCCAATTCAACAGAGGATACACACCGAGACTCGAGGGGAC	2005
Db	2379	GGACGTCTTTTTTTTAATGCTCCCAATTCAACAGAGGATACACACCGAGACTCGAGGGGAC	2438
Qy	2006	ACGATAGAGCACCAAGGCCCCACCTTTGTCCCTTGAGAAATTCAGGGTACAGCTCACCTAG	2065
Db	2439	ACGATAGAGCACCAAGGCCCCACCTTTGTCCCTTGAGAAATTCAGGGTACAGCTCACCTAG	2498
Qy	2066	AACCAAGGCTGCCCTCTACTGTGCTCTGTTCTCCACAACTTATTTAAGACCTACACTGGGTGC	2125
Db	2499	AACCAAGGCTGCCCTCTACTGTGCTCTGTTCTCCACAACTTATTTAAGACCTACACTGGGTGC	2558
Qy	2126	TGGGTTCACTGTGCTCTAGAGAAACCAAGAGGCTCCCACTCTGAGCTCTCTGCCGCCCT	2185
Db	2559	TGGGTTCACTGTGCTCTAGAGAAACCAAGAGGCTCCCACTCTGAGCTCTCTGCCGCCCT	2618
Qy	2186	GCTGCCCAACCACTTTCTGCAACACAGGCTGGGGAGGCGGGAGAGACAGCTTGGACCC	2245
Db	2619	GCTGCCCAACCACTTTCTGCAACACAGGCTGGGGAGGCGGGAGAGACAGCTTGGACCC	2678
Qy	2246	CAGAACTGAGCCTGGGAGGGATCCGACAGAAAAGCTCAGGGCGGGTCTCTCTCTGTGC	2304
Db	2679	CAGAACTGAGCCTGGGAGGGATCCGACAGAAAAGCTCAGGGCGGGTCTCTCTCTGTGC	2738
Qy	2305	CCGGGATTGGGCTATGCTGGGTACCAACATGTACTCAGGCATGTGGGTTTGAACCAT	2364
Db	2739	CCGGGATTGGGCTATGCTGGGTACCAACATGTACTCAGGCATGTGGGTTTGAACCAT	2798
Qy	2365	AAACCAAAAGCCCTTGTCATCAGCTTTTAAACAAGTATATTTGATTTTAACTCTCAA	2424
Db	2799	AAACCAAAAGCCCTTGTCATCAGCTTTTAAACAAGTATATTTGATTTTAACTCTCAA	2855
Qy	2425	ACATATTGAAGTTTATGGGCCCTTAAGAACCTTAGTATCTTCTATTTGGCTTTTCTAG	2484
Db	2859	ACATATTGAAGTTTATGGGCCCTTAAGAACCTTAGTATCTTCTATTTGGCTTTTCTAG	2918
Qy	2485	GTTTCAGAGAGGTTAAGTAACTTCTCTCAGG-TCACACAGAACTCTGTGGGTGGAGAG	2544
Db	2919	GTTTCAGAGAGGTTAAGTAACTTCTCTCAGGTTCAACAGAACTCTGTGGGTGGAGAG	2978

QY	2544	CAAGCTACCGCTGGGGCATTTAGATACATACACAGATGGTCCCTCTTATGATGCTTGGCC	2603
Db	2979	CAACCTTGCACATGGGGCATTAAGTACATACACAGATGGTCCCTCTTATGATGCTTGGCC	3038
QY	2604	CCTGGGGCCATTCAGGGCTTTGGGACATCTGTCTCAACCCTCTCCCTAATCAGTCTGT	2663
Db	3039	CTTGGGGCCATTCAGGGCTTTGGGACATCTGTCTCAACCCTCTCCCTAATCAGTCTGT	3098
QY	2664	GAGGCTCCCTGTATATTTGTGTATACACCATGGCCATGTATATACAAATACACAGATGT	2723
Db	3099	GAGGCTCCCTGTATATTTGTGTATACACCATGGCCATGTATATACAAATACACAGATGT	3158
QY	2724	ACACACAGATGTACACATGCTCCAGCCCGACGCTCTGTATACCTGCACCTGCACCCGACCC	2783
Db	3159	ACACACAGATGTACACATGCTCCAGCCCGACGCTCTGTATACCTGCACCTGCACCCGACCC	3218
QY	2784	TTGGCCCTGCTGCTGCTGTGTGTCTTAAGCAGACAGCTTCAACCTCTGCTCTCTCTCCCTTC	2843
Db	3219	TTGGCCCTGCTGCTGCTGTGTGTCTTAAGCAGACAGCTTCAACCTCTGCTCTCTCTCCCTTC	3278
QY	2844	CCCAACCACTGCTCGAGGCTTCTAGAGCAGACACAGTACCTTGGCTGCACCGGTGTGGGC	2903
Db	3279	CCCAACCACTGCTCGAGGCTTCTAGAGCAGACACAGTACCTTGGCTGCACCGGTGTGGGC	3338
QY	2904	CCGCTCTACCCAGGACAGGCCCGGCCACATGATCTCCGTGTACACTATCAATTAAG	2963
Db	3339	CCGCTCTACCCAGGACAGGCCCGGCCACATGATCTCCGTGTACACTATCAATTAAG	3398
QY	2964	TGGGTTTGTACAAAAAATAAAAA 2989	
Db	3399	TGGGTTTGTACAAAAAATAAAAA 3424	

	RESULT 5				
	AX393244				
LOCUS	AX393244	2832 bp	DNA	linear	PAT 23-MAR-2002
DEFINITION	Sequence 174 from Patent WO0210217.				
ACCESSION	AX393244				
VERSION	AX393244.1	GI:19701294			
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1				
AUTHORS	St Croix,B., Kinzler,K.W. and Vogelstein,B.				
TITLE	Endothelial cell expression patterns				
JOURNAL	Patent: WO 0210217-A 174 07-FEB-2002;				
	The Johns Hopkins University (US)				
FEATURES	Location/Qualifiers				
source	1..2832				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
	Query Match 91.1%; Score 2806.4; DB 6; Length 2832;				
	Best Local Similarity 100.0%; Pred.No. 0;				
	Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0				
QY	199 CAGCCCCGAGCCATGTAAGAAGCTTGTTCACAGCGGAACGTGCACGCTCAGTGTGCCGGC 258				
Db	25 CAGCCCCGAGCCATGTAAGAAGCTTGTTCACAGCGGAACGTGCACGCTCAGTGTGCCGGC 84				
QY	259 AAAAATCATACCGATGTTGTGCTGGTGCTCTTCGGGTGGCAAGACTTCATCGTG 318				
Db	85 AAAAATCATACCGATGTTGTGCTGGTGCTCTTCGGGTGGCAAGACTTCATCGTG 144				
QY	319 TCTCGTTCTCTCAATAGCCGCTTTTAGACAACAGTACACACCACCATGAGAGATTTCAC 378				
Db	145 TCTCGTTCTCTCAATAGCCGCTTTTAGACAACAGTACACACCACCATGAGAGATTTCAC 204				

QY 379 GGTAAAGTATACAAATCCGCGGCGACAGATGACAGCTCGACATCTCGAATACCTCTGGC 438  
 DB 205 CGTAAGGATATACAAATCCGCGGCGACAGATGACAGCTCGACATCTCGAATACCTCTGGC 264  
 QY 439 AACCAACCCCTTCCCGCATGCGAGGCTGTCATCTCAACAGGGAATGCTTCACTCTG 498  
 DB 265 AACCAACCCCTTCCCGCATGCGAGGCTGTCATCTCAACAGGGAATGCTTCACTCTG 324  
 QY 499 GTGTTCAACCTGGAATAAACCGGGAATGCTTCAATGAGTCAAGGCTTCAAGAGATG 558  
 DB 325 GTGTCAACCTGGAATAAACCGGGAATGCTTCAATGAGTCAAGGCTTCAAGAGATG 384  
 QY 559 CTGAGAGTCAAGTCTGCTGCTGAAGAAACAAGACAGAGAGCGCGAGAGCTGCGCATG 618  
 DB 385 CTGAGAGTCAAGTCTGCTGCTGAAGAAACAAGAGAGCGCGAGAGCTGCGCATG 444  
 QY 619 ATCTGTGGCAACAGAACGACCAACGCGAGCTGTGCGCGAGAGTCCCAACCAACGAGCC 678  
 DB 445 ATCTGTGGCAACAGAACGACCAACGCGAGCTGTGCGCGAGAGTCCCAACCAACGAGCC 504  
 QY 679 GAGCTGCTGTGTGGGGGAGAGAACTCCGCTTCTTCAAGTGTGCGCGCAAGAGAAC 738  
 DB 505 GAGCTGCTGTGTGGGGGAGAGAACTCCGCTTCTTCAAGTGTGCGCGCAAGAGAAC 564  
 QY 739 ACCAAGCTGAGACGAGATGTTCTACGTGCTTCAAGCATGAGCCAGCTGCAACAGAGATG 798  
 DB 565 ACCAAGCTGAGACGAGATGTTCTACGTGCTTCAAGCATGAGCCAGCTGCAACAGAGATG 624  
 QY 799 AGCCCGCGCTGCAATCGCAAGATCTCCGTGACGATGAGCGCTTCAACCCCAAGCC 858  
 DB 625 AGCCCGCGCTGCAATCGCAAGATCTCCGTGACGATGAGCGCTTCAACCCCAAGCC 684  
 QY 859 TTTCGACATGCGCGCGCTGCAAGAGATGAGACCGCTTATGAGATGATGCTGCGCGCTG 918  
 DB 685 TTTCGACATGCGCGCGCTGCAAGAGATGAGACCGCTTATGAGATGATGCTGCGCGCTG 744  
 QY 919 CGCCCAAGCTCAACAGTGAACCTCAAGTACATCAAGGCAAGGCTCTTCCGGAAGAGCCAG 978  
 DB 745 CGCCCAAGCTCAACAGTGAACCTCAAGTACATCAAGGCAAGGCTCTTCCGGAAGAGCCAG 804  
 QY 979 GCCCGTGAAGAGGAGCAAGTGAACCAATCCAGTGAAGGAGATGCTGAGGCGGAGCTTGGC 1038  
 DB 805 GCCCGTGAAGAGGAGCAAGTGAACCAATCCAGTGAAGGAGATGCTGAGGCGGAGCTTGGC 864  
 QY 1039 CAGTGCCTTCAAGGAGGAGGCGCCCAAGATGCCCCATCTGCCCATCCAGAGCCCC 1098  
 DB 865 CAGTGCCTTCAAGGAGGAGGCGCCCAAGATGCCCCATCTGCCCATCCAGAGCCCC 924  
 QY 1099 GGCAGACAGTCTTGTTCACAGACCTTAAGGACACAGCTGGAAGGCCCGGAGCGCTGGCCTC 1158  
 DB 925 GGCAGACAGTCTTGTTCACAGACCTTAAGGACACAGCTGGAAGGCCCGGAGCGCTGGCCTC 984  
 QY 1159 CGCAATTCGATCTGCTTCTCAAGCTTTCGATGATGCTGCTGCTCAAGCTCTTGGTG 1218  
 DB 985 CGCAATTCGATCTGCTTCTCAAGCTTTCGATGATGCTGCTGCTCAAGCTCTTGGTG 1044  
 QY 1219 GTTTCATATCTCTGTGGAAGAGACATCTGTGACGCTCAAGAGTTGAGAGACTCA 1278  
 DB 1045 GTTTCATATCTCTGTGGAAGAGACATCTGTGACGCTCAAGAGTTGAGAGACTCA 1104  
 QY 1279 AGTTACACCTTCTCTCTGAGGAGTGAAGAAATGTATGACAGAGGAGTGAAGATTGC 1338  
 DB 1105 AGTTACACCTTCTCTCTGAGGAGTGAAGAAATGTATGAGAGGAGTGAAGATTGC 1164  
 QY 1339 TGCGTCATATGAGAGCTCTGAGGACAAAGCTCAAGATGAAAGACACAGAGAGCCAGAT 1398  
 DB 1165 TGCGTCATATGAGAGCTCTGAGGACAAAGCTCAAGATGAAAGACACAGAGAGCCAGAT 1224  
 QY 1399 GAGAAAGTCTCTCTCTCTGAGCATACACCAAGCTGTGTTGGTGTGAGCTGGAGGA 1458  
 DB 1225 GAGAAAGTCTCTCTCTCTGAGCATACACCAAGCTGTGTTGGTGTGAGCTGGAGGA 1284  
 QY 1459 ACTTCTCTCCAGCCCTGCAACTTACGCTCTGAGTTCAGCTGCTTGCACCCCTCTCC 1518

DB 1285 ACTTCTCTCCAGCCCTGCAACTTACGCTCTGAGTTCAGCTGCTGCAACCCCTCTCC 1344  
 QY 1519 ACCCCAGACACACACAAAGTTGGCCCGCAGCTGCGCTGACATTTAGCCAGTGAATCT 1578  
 DB 1345 ACCCCAGACACACACAAAGTTGGCCCGCAGCTGCGCTGACATTTAGCCAGTGAATCT 1404  
 QY 1579 GTGTCTGAAGGGGCGTGGCCACACCTCTGACACAGGCCACACATTTAGACACAGGCCA 1638  
 DB 1405 GTGTCTGAAGGGGCGTGGCCACACCTCTGACACAGGCCACACATTTAGACACAGGCCA 1464  
 QY 1639 CCTCTGACCGCGTTCCTCAAGCTCTCTCTCTAGAGCTCCCGCGCAGACAGTGTGCTTT 1698  
 DB 1465 CCTCTGACCGCGTTCCTCAAGCTCTCTCTCTAGAGCTCCCGCGCAGACAGTGTGCTTT 1524  
 QY 1699 GTTGTGTTGACGCTGTTTCTGTGTCATGATATAGTATGATAAATGAAATCATTTGACTG 1758  
 DB 1525 GTTGTGTTGACGCTGTTTCTGTGTCATGATATAGTATGATAAATGAAATCATTTGACTG 1584  
 QY 1759 TAAAGCCTAGTGAATCCCTCTTGGCCAGGCCCTCAACCCAGTTCAATCCAGCCCTCC 1818  
 DB 1585 TAAAGCCTAGTGAATCCCTCTTGGCCAGGCCCTCAACCCAGTTCAATCCAGCCCTCC 1644  
 QY 1819 ACCCGGAGCGCCTTCTCTCTCTGCTCCCAACAGGGTTTCCGTGGCCTGTTTGAAGCTAG 1878  
 DB 1645 ACCCGGAGCGCCTTCTCTCTCTGCTCCCAACAGGGTTTCCGTGGCCTGTTTGAAGCTAG 1704  
 QY 1879 ACATTTAGCTCGCATTTAGCTCAACGTTTACAGACATTTGACAAAGCGTGGGTGGG 1938  
 DB 1705 ACATTTAGCTCGCATTTAGCTCAACGTTTACAGACATTTGACAAAGCGTGGGTGGG 1764  
 QY 1939 CAGGCGCAGAACTGCTTTTAAATGCTCCCATTTCAAGAGATACCAACGAGACTGG 1998  
 DB 1765 CAGGCGCAGAACTGCTTTTAAATGCTCCCATTTCAAGAGATACCAACGAGACTGG 1824  
 QY 1999 AGGGGACACGATGAGACACAGGCCCACTTGTCTCCCTAGCAATTTCAAGGTTACAGCTC 2058  
 DB 1825 AGGGGACACGATGAGACACAGGCCCACTTGTCTCCCTAGCAATTTCAAGGTTACAGCTC 1884  
 QY 2059 CACCTAGAACAGGCTGCTCTTACTGTGCTGTTTCTCAAGCATTTATTAAGACCTTAC 2118  
 DB 1885 CACCTAGAACAGGCTGCTCTTACTGTGCTGTTTCTCAAGCATTTATTAAGACCTTAC 1944  
 QY 2119 TGGGTGCTGGGTTCACGTGTGCTTGAAGAAACCAAGAGGTCCTCCAGTCTGGCTTGGC 2178  
 DB 1945 TGGGTGCTGGGTTCACGTGTGCTTGAAGAAACCAAGAGGTCCTCCAGTCTGGCTTGGC 2004  
 QY 2179 CGCCCTGCTGCGCCCAACACTTCTGACACACAGCGGTGGGAGGCGGGAGAGACAGC 2238  
 DB 2005 CGCCCTGCTGCGCCCAACACTTCTGACACACAGCGGTGGGAGGCGGGAGAGACAGC 2064  
 QY 2239 TGGGACCCAGAACTGAGCTTGGAGGAGATCCGACAGAAAGCTCAGGGCGGGTCTTCC 2298  
 DB 2065 TGGGACCCAGAACTGAGCTTGGAGGAGATCCGACAGAAAGCTCAGGGCGGGTCTTCC 2124  
 QY 2299 TTTGCCCCGGGATTTGGGCTATGCTGGGTACCAACATGATCTCAGGCAATGGTGGTTTGA 2358  
 DB 2125 TTTGCCCCGGGATTTGGGCTATGCTGGGTACCAACATGATCTCAGGCAATGGTGGTTTGA 2184  
 QY 2359 ACCCATTAACCAAGAGCCCTTGTATCATGAGCTTTAAACAGTATTTTAAATCT 2418  
 DB 2185 ACCCATTAACCAAGAGCCCTTGTATCATGAGCTTTAAACAGTATTTTAAATCT 2244  
 QY 2419 CTCTAAACATATTTGAAGTTTGAAGGCTTAAGGAACCTTATGATCTTATTTGGTCTT 2478  
 DB 2245 CTCTAAACATATTTGAAGTTTGAAGGCTTAAGGAACCTTATGATCTTATTTGGTCTT 2304  
 QY 2479 TCTGAGGTTCAAGAGGTTAAGTATCTTCTCCAGGTTCACACAGCAAGTCTGTGGGTGGC 2538  
 DB 2205 TCTGAGGTTCAAGAGGTTAAGTATCTTCTCCAGGTTCACACAGCAAGTCTGTGGGTGGC 2364  
 QY 2539 AGAAGCAAGCTAAGGCTGGGACCTTCAATACATACAGATGTGCTCTCTGATGCT 2598



Db 2365 AGAAGCAGCTAGCGCTGGGCAATTCAGTACATACACAGATGTCTCCCTCTTGAATGCT 2424  
 QY 2599 TGGCCCCCTGGGCGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTCTCCCTAGATCAG 2658  
 Db 2425 TGGCCCCCTGGGCGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTCTCCCTAGATCAG 2484  
 QY 2659 TCTGTAGGAGTCTCTGTAGATATTTGTGTACACCATGATATATACAGTACACACA 2718  
 Db 2485 TCTGTAGGAGTCTCTGTAGATATTTGTGTACACCATGATATATACAGTACACACA 2544  
 QY 2719 GATGTACACACAGATGTACACATGTCTCAGGCCCGAGCTGTGATCTGACATCTGACACCC 2778  
 Db 2545 GATGTACACACAGATGTACACATGTCTCAGGCCCGAGCTGTGATCTGACATCTGACACCC 2604  
 QY 2779 CAGCCTTGGCCCTCTGCTGCTGTGTGTCAAAAGCAGACAGCTCCAACTCTGCTGTCTC 2838  
 Db 2605 CAGCCTTGGCCCTCTGCTGCTGTGTGTCAAAAGCAGACAGCTCCAACTCTGCTGTCTC 2664  
 QY 2839 CTTTCCCAACCCACATGCTGCTGAGCTTTCTGAGACACAGTACCTTGGCTGACCCGGTGT 2898  
 Db 2665 CTTTCCCAACCCACATGCTGCTGAGACACAGTACCTTGGCTGACCCGGTGT 2724  
 QY 2899 GTGGCCCGCTCTCACCACAGGACAGCCCGCACCAGTGTCTCCGTGTACATCTATCAAT 2958  
 Db 2725 GTGGCCCGCTCTCACCACAGGACAGCCCGCACCAGTGTCTCCGTGTACATCTATCAAT 2784  
 QY 2959 AAAAGTGGCTTTGTGTACAAAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3006  
 Db 2785 AAAAGTGGCTTTGTGTACAAAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2832

RESULT 6  
 AF279143 2832 bp mRNA linear PRI 23-AUG-2000  
 LOCUS Homo sapiens tumor endothelial marker 2 (TEM2) mRNA, complete cds.  
 DEFINITION AF279143  
 ACCESSION AF279143.1 GI:9857401  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Homo sapiens Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 St. Croix, B., Rago, C., Velculescu, V., Traverson, G., Romans, K. B.,  
 Montgomery, E., Lai, A., Riggin, G. D., Lengauer, C., Vogelstein, B. and  
 Kinzler, K. W.  
 Genes expressed in human tumor endothelium  
 Science 289 (5482), 1197-1202 (2000)  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED 10947988  
 2 (bases 1 to 2832)  
 St. Croix, B., Vogelstein, B. and Kinzler, K. W.  
 Direct Submission  
 Submitted (16-JUN-2000) Johns Hopkins Oncology Center, Johns  
 Hopkins University, 1650 Orleans Street, Baltimore, MD 21231, USA  
 FEATURES  
 source  
 1..2832  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 /map="12q13.1"  
 1..2832  
 /gene="TEM2"  
 1..837  
 /gene="TEM2"  
 /note="TEM2 protein"  
 /codon\_start=1  
 /product="tumor endothelial marker 2"  
 /protein\_id="AA00868.1"  
 /db\_xref="GI:9857402"  
 /translation="MPASLALLOPRAMKTLSSGNTLSVPANRKYVIRMYVTLGASRYGK  
 SSIVSRFLNGRFEDQYPTPTIEDFKRYNIRGDMYQOLDIDTSGNHPFAMRLSILT

ORIGIN  
 Query Match 91.1%; Score 2806.4; DB 9; Length 2832;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 10"  
 GGVFLIVFSLDNRSEFDEYKRLQKOLLEYSCLKRKTEAAELPMVIGKNRNGELC  
 ROVPTTEALIVSGDENCAVEFSAKKNVNVEMFVLESMALPHMRKISV  
 QYGDAPHRPFPCKMRKYKENDAYGMSPFARRSVNSDLKYLIAKYLIREQARERDXT  
 199 CAGCCCGAGCCATGATGAAGACTTTGTGTCAGCGGGAATGACAGCTGAGTGTCCCGCC 258  
 Db 25 CAGCCCGAGCCATGATGAAGACTTTGTGTCAGCGGGAATGACAGCTGAGTGTCCCGCC 84  
 QY 259 AAAAATCTCAATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 318  
 Db 85 AAAAATCTCAATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 144  
 QY 319 TCTGCTTCTCAATGAGCCGCTTTGAGGACAGTACACACCCATCGAGACTTCAC 378  
 Db 145 TCTGCTTCTCAATGAGCCGCTTTGAGGACAGTACACACCCATCGAGACTTCAC 204  
 QY 379 CGTAAAGTATACAAATCCCGCGGACATGTATACAGTGTGATCCTGTGATCTGAGC 438  
 Db 205 CGTAAAGTATACAAATCCCGCGGACATGTATACAGTGTGATCCTGTGATCTGAGC 264  
 QY 439 AACCAACCCCTTCCCGCGCATGCGGAGCTGTCTCATCTCTCAGAGGAGTCTTCATCTTG 498  
 Db 265 AACCAACCCCTTCCCGCGCATGCGGAGCTGTCTCATCTCTCAGAGGAGTCTTCATCTTG 324  
 QY 499 GTGTTCAAGCTGTGATTAACCGGAGTCTTGTGATGATGATGATGATGATGATGATGATG 558  
 Db 325 GTGTTCAAGCTGTGATTAACCGGAGTCTTGTGATGATGATGATGATGATGATGATGATG 384  
 QY 559 CTGAGAGTCAAGTCTCTCTGTAAGAACAAAGAGAGGCGGAGCTGAGCCATGATGTC 618  
 Db 385 CTGAGAGTCAAGTCTCTCTGTAAGAACAAAGAGAGGCGGAGCTGAGCCATGATGTC 444  
 QY 619 ATCTGTGCAACAGAAAGACCAACGAGAGTGTGCGGACAGTGTGCGGACCAACCGAGGCT 678  
 Db 445 ATCTGTGCAACAGAAAGACCAACGAGAGTGTGCGGACAGTGTGCGGACCAACCGAGGCT 504  
 QY 679 GAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 738  
 Db 505 GAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 564  
 QY 739 ACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 798  
 Db 565 ACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 624  
 QY 799 AGCCCGGCGCTGATTCGCAAGATCTCCGTGACGATGACGATGACGATGACGATGACGATG 858  
 Db 625 AGCCCGGCGCTGATTCGCAAGATCTCCGTGACGATGACGATGACGATGACGATGACGATG 684  
 QY 859 TTCTGCAATGCGCGCGGTGTAAGAGATGAGAGCCGATGATGATGATGATGATGATGATG 918  
 Db 685 TTCTGCAATGCGCGCGGTGTAAGAGATGAGAGCCGATGATGATGATGATGATGATGATG 744  
 QY 919 CGCCCGAGGTCACAGTACCTCAAGTACATCAAGGCAAGGCTCTTGGGAAAGCCAG 978  
 Db 745 CGCCCGAGGTCACAGTACCTCAAGTACATCAAGGCAAGGCTCTTGGGAAAGCCAG 804  
 QY 979 GCCCGTGAAGAGGACCAATGTGATGATGATGATGATGATGATGATGATGATGATGATG 1038  
 Db 805 GCCCGTGAAGAGGACCAATGTGATGATGATGATGATGATGATGATGATGATGATGATG 864  
 QY 1039 CAGTGTCTTCAAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1098  
 Db 865 CAGTGTCTTCAAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 924  
 QY 1099 GCGAGAGTCTTGTTCACAGACCTTAAGGACACAGACTGAGAGGCGCCGAGGCTGAGCTC 1158  
 Db 925 GCGAGAGTCTTGTTCACAGACCTTAAGGACACAGACTGAGAGGCGCCGAGGCTGAGCTC 984



QY 1159 CGCATTGCTGCTCTCTCAAGCTTCTGAGTCCGCTGTCCAGAGCTCTTGGTG 1218  
 DB 985 CGCATTGCTGCTCTCTCAAGCTTCTGAGTCCGCTGTCCAGAGCTCTTGGTG 1044  
 QY 1219 GTTTCATCTCTCTGTGGAGAGACATCTGTGAGCTCAAGAGTTCAGAGACTCA 1278  
 DB 1045 GTTTCATCTCTCTGTGGAGAGACATCTGTGAGCTCAAGAGTTCAGAGACTCA 1104  
 QY 1279 AGTTACACTTCTCTCTGTGGAGAGATTTGATGTCAGAGGGGTGAGATTCG 1338  
 DB 1105 AGTTACACTTCTCTCTGTGGAGAGATTTGATGTCAGAGGGGTGAGATTCG 1164  
 QY 1339 TGCCGATATGAGAGCTCTGTGGAGAGAGCTCAAGAGTTCAGAGAGCTCAAGAT 1398  
 DB 1165 TGCCGATATGAGAGCTCTGTGGAGAGAGCTCAAGAGTTCAGAGAGCTCAAGAT 1224  
 QY 1399 GAGAAAGGTCTCTCTCTGTGGAGATTAACCCAGCTTGTGGGTGGAGCTGGAGA 1458  
 DB 1225 GAGAAAGGTCTCTCTCTGTGGAGATTAACCCAGCTTGTGGGTGGAGCTGGAGA 1284  
 QY 1459 ACTTCTCTCCAGGCTCTGCACTCTTAAGCTTGTGGGTGAGCTCTGACCTCTCC 1518  
 DB 1285 ACTTCTCTCCAGGCTCTGCACTCTTAAGCTTGTGGGTGAGCTCTGACCTCTCC 1344  
 QY 1519 ACCCCAGAGACACACAAAGTTGGCCCGAGCTGGCTGACATTTAGAGCTGGAGCT 1578  
 DB 1345 ACCCCAGAGACACACAAAGTTGGCCCGAGCTGGCTGACATTTAGAGCTGGAGCT 1404  
 QY 1579 GTGTCTGAAGGGGGGTGGGCAACCTCTAGACAGCCCACTTAAGACAGCCCA 1638  
 DB 1405 GTGTCTGAAGGGGGGTGGGCAACCTCTAGACAGCCCACTTAAGACAGCCCA 1464  
 QY 1639 CCTCTGACCGGCTCTCTGAGCTCTCTCTAGTCCCTCCGCGGACAGTTGTCTTT 1698  
 DB 1465 CCTCTGACCGGCTCTCTGAGCTCTCTCTAGTCCCTCCGCGGACAGTTGTCTTT 1524  
 QY 1699 GTTGTGATGAGAGCTTTGTGTGATGATGATGATGATGATGATGATGATGATG 1758  
 DB 1535 GTTGTGATGAGAGCTTTGTGTGATGATGATGATGATGATGATGATGATGATG 1584  
 QY 1759 TAAAGAGCTAGTGAATCTCTCTGTGGAGAGCTCAAGCTTGAATGATGATGATG 1818  
 DB 1585 TAAAGAGCTAGTGAATCTCTCTGTGGAGAGCTCAAGCTTGAATGATGATGATG 1644  
 QY 1819 ACCGGGAGAGCT 1878  
 DB 1645 ACCGGGAGAGCT 1704  
 QY 1879 ACATTTGATCT 1938  
 DB 1705 ACATTTGATCT 1764  
 QY 1939 CAGGAGAGAGCT 1998  
 DB 1765 CAGGAGAGAGCT 1824  
 QY 1999 AGGGAGACAGTGAAGACAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2058  
 DB 1825 AGGGAGACAGTGAAGACAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1884  
 QY 2059 CACCTAGAGACAGGCT 2118  
 DB 1885 CACCTAGAGACAGGCT 1944  
 QY 2119 TGGGTGCTGAGTCACTGTGTCTTGAAGAACCAAGAGGTCTCCAGTCTGTGCTCTCC 2178  
 DB 1945 TGGGTGCTGAGTCACTGTGTCTTGAAGAACCAAGAGGTCTCCAGTCTGTGCTCTCC 2004  
 QY 2179 CGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2238  
 DB 2005 CGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2064

QY 2239 TGGAGCCAGAACTGAGCTGTGGAGAGATCCGACAGAAAGCTGAGGGGCTCTTCC 2298  
 DB 2065 TGGAGCCAGAACTGAGCTGTGGAGAGATCCGACAGAAAGCTGAGGGGCTCTTCC 2124  
 QY 2299 TTGTGCCCCGAGATTTGGCTATGCTGGGTACCAACATGTAATCAGGATGTGGTTTGA 2358  
 DB 2125 TTGTGCCCCGAGATTTGGCTATGCTGGGTACCAACATGTAATCAGGATGTGGTTTGA 2184  
 QY 2359 ACCCATTAACCAAGAGGCTCTGTATGATGATGATGATGATGATGATGATGATG 2418  
 DB 2185 ACCCATTAACCAAGAGGCTCTGTATGATGATGATGATGATGATGATGATGATG 2244  
 QY 2419 CTCTAACAATTAAGTTTAAAGGCTTAAAGAACTTAAAGAACTTAAAGAACTTAAAG 2478  
 DB 2245 CTCTAACAATTAAGTTTAAAGGCTTAAAGAACTTAAAGAACTTAAAGAACTTAAAG 2304  
 QY 2479 TCTGAGGTTCAAGAGAGGTAAATTAATCTCTCAAGTCAACAGCAAGCTGTGGGTGC 2538  
 DB 2305 TCTGAGGTTCAAGAGAGGTAAATTAATCTCTCAAGTCAACAGCAAGCTGTGGGTGC 2364  
 QY 2539 AGAAGCAAGCTAAGGCTGTGGGATTCAGTACATTAACAGATGCTCTCTTGAATGCT 2598  
 DB 2365 AGAAGCAAGCTAAGGCTGTGGGATTCAGTACATTAACAGATGCTCTCTTGAATGCT 2424  
 QY 2599 TGAGCCCTGTGGGCTTCAAGGCTTGGGACATCTTGTCTCAACCTCTCTCTAGATGCT 2658  
 DB 2425 TGAGCCCTGTGGGCTTCAAGGCTTGGGACATCTTGTCTCAACCTCTCTCTAGATGCT 2484  
 QY 2659 TCTGTGAGAGGTCTCTGTATGATGATGATGATGATGATGATGATGATGATGATG 2718  
 DB 2485 TCTGTGAGAGGTCTCTGTATGATGATGATGATGATGATGATGATGATGATGATG 2544  
 QY 2719 GATGTACACACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2778  
 DB 2545 GATGTACACACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2604  
 QY 2779 CAGCCTTGGGCTCTGCTGCTGTGTCTCAAGAGAGCTTCAACCTCTCTCTGCTC 2838  
 DB 2605 CAGCCTTGGGCTCTGCTGCTGTGTCTCAAGAGAGCTTCAACCTCTCTCTGCTC 2664  
 QY 2839 CTTTCCCAACCACTGCTCTGAGCTTGTAGACAGACAGTACTTGGCTGACCGGTGT 2898  
 DB 2665 CTTTCCCAACCACTGCTCTGAGCTTGTAGACAGACAGTACTTGGCTGACCGGTGT 2724  
 QY 2899 GTGGCCGCTCTCAACCAAGGACAGGCTTCAACAGATGATGATGATGATGATGAT 2958  
 DB 2725 GTGGCCGCTCTCAACCAAGGACAGGCTTCAACAGATGATGATGATGATGATGAT 2784  
 QY 2959 AAAAGTGGGTTGTATCAAAAAAAGGCTTCAAAAAAAGGCTTCAAAAAAAGGCTT 3006  
 DB 2785 AAAAGTGGGTTGTATCAAAAAAAGGCTTCAAAAAAAGGCTTCAAAAAAAGGCTT 2832

RESULT 7  
 HSM803172  
 LOCUS HSM803172 2699 bp mRNA linear PRI 10-JUL-2002  
 DEFINITION Homo sapiens mRNA; cDNA DKFZp761D0217 (from clone DKFZp761D0217).  
 ACCESSION AL631936  
 VERSION AL631936.1 GI:21732458  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2699)  
 AUTHORS Ansoorge, W., Wilkner, U., Mewes, H. W., Weil, B. and Wiemann, S.  
 TITLE Submitted (09-JUL-2002) 1. D-85764 Neuberberg, GERMANY  
 JOURNAL Clome from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 COMMENT Heidelberg/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.

This clone (DKFZp761D0217) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cDNA/>.

## FEATURES

## source

1..2699  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="RZPD:DKFZp761D0217"  
/db\_xref="taxon:9606"  
/map="22q13.1"  
/clone="DKFZp761D0217"  
/cissue\_type="amygdala"  
/clone\_lib="761 (synonym: hamy2). Vector pSport1; host  
DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
2660..2665  
polyA\_signal  
polyA\_site  
2677

## ORIGIN

Query Match 86.0%; Score 2651.6; DB 9; Length 2699;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 2690; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

294 TCGGGTGGGCAAGAGCTCCATGTCGTCTTCGCTTCTCAATGAGCCGCTTTAGAGCACTA 353  
1 TCGAGTGGGCAAGAGCTCCATGTCGTCTTCGCTTCTCAATGAGCCGCTTTAGAGCACTA 60  
354 CACACCCACATCGAGAGATTTCACCGTAAAGTATACATCCGCGGGAATGATACCA 413  
61 CACACCCACATCGAGAGATTTCACCGTAAAGTATACATCCGCGGGAATGATACCA 120  
414 GCTCGACATCTGATACCTCTGCAACCAACCCCTTCCCGGCATGCGAGGCTGTCAAT 473  
121 GCTCGACATCTGATACCTCTGCAACCAACCCCTTCCCGGCATGCGAGGCTGTCAAT 180  
474 CCTCAGAGGGAGTGTCTTCACTCTGATGTTCAAGCTGTGATTAACCGGAGATCTTGTGATTA 533  
181 CCTCAGAGGGAGTGTCTTCACTCTGATGTTCAAGCTGTGATTAACCGGAGATCTTGTGATTA 240  
534 GGTTCAGAGCCCTTCAAGACAGATCCTGAGAGTCAAGTCTGCTGAAGAACAGACCA 593  
241 GGTTCAGAGCCCTTCAAGACAGATCCTGAGAGTCAAGTCTGCTGAAGAACAGACCA 300  
594 GAGAGCGGCGAGAGCTCCATGTCATCTGTGCAACAAAGACGACGCGAGCTGTG 653  
301 GAGAGCGGCGAGAGCTCCATGTCATCTGTGCAACAAAGACGACGCGAGCTGTG 360  
654 CCGCCAGGTGCCACCAACCGAGGCGGAGCTGTGTGGGGAAGAACTGCGCTTA 713  
361 CCGCCAGGTGCCACCAACCGAGGCGGAGCTGTGTGGGGAAGAACTGCGCTTA 420  
714 CTTGAGGTGTGGGCAAGAAACCAACGATGAGAGATGTTCTAGTGTCTTCAAG 773  
421 CTTGAGGTGTGGGCAAGAAACCAACGATGAGAGATGTTCTAGTGTCTTCAAG 480  
774 CATGCGCAAGCTGCCACAGAGATGAGCCCGCTTGCATGCAAGATCTCCGTGACATA 833  
481 CATGCGCAAGCTGCCACAGAGATGAGCCCGCTTGCATGCAAGATCTCCGTGACATA 540  
834 CGGTGAGCTTCCACCCAGGCGCTTGTGATGAGCCCGGCTGAAGAGATGAGACGCTTA 893  
541 CGGTGAGCTTCCACCCAGGCGCTTGTGATGAGCCCGGCTGAAGAGATGAGACGCTTA 600  
894 TGGCATGATCTCGCCCTTCCGCGCGCCGACAGCTCAAGTGAACCTCAAGTACATCA 953  
601 TGGCATGATCTCGCCCTTCCGCGCGCCGACAGCTCAAGTGAACCTCAAGTACATCA 660  
954 GGCCTAAGTCTTTCGGAAGGCGAGGCGCTGAGAGGGAACAATGACCATTCAGTAGAC 1013  
661 GGCCTAAGTCTTTCGGAAGGCGAGGCGCTGAGAGGGAACAATGACCATTCAGTAGAC 719

1014 GAGGATGCTGGGGCGGGGCTTGCCAGATGCTTCAGGAGAGGTGCGCCAGATGCCACT 1073  
720 GAGGATGCTGGGGCGGGGCTTGCGAGTCTTCAGGAGAGGTGCGCCAGATGCCACT 777  
1074 GTGGGATCTCCCAACGAGGCGCCGAGCAGTCTTGTTCACAGACTTTAGGACACCA 1133  
778 GTGGGATCTCCCAACGAGGCGCCGAGCAGTCTTGTTCACAGACTTTAGGACACCA 837  
1134 CTGAGAGGCGCCGCGGCGCTGAGCTCCGACATTCGTCTGCTTCTCAAGCTTCTCGAG 1193  
838 CTGAGAGGCGCCGCGGCGCTGAGCTCCGACATTCGTCTGCTTCTCAAGCTTCTCGAG 897  
1194 TCGGCTGTGCAAGAGCTCTTGGGTGTTTCACTCTCTGTGAGAGACATCTCTGCA 1253  
898 TCCCTGTGCAAGAGCTCTTGGGTGTTTCACTCTCTGTGAGAGACATCTCTGCA 957  
1254 GCTTCAGAGATTAGGCAAGAGCTCAAGTTACACTTCTCTCTGAGGTTGAAAGATG 1313  
958 GCTTCAGAGATTAGGCAAGAGCTCAAGTTACACTTCTCTCTGAGGTTGAAAGATG 1017  
1314 TTGATGCCAAGAGGAGTATGAGTCTGCTCATATGAGAGCTCTGAGAGCAAGCTCAG 1373  
1018 TTGATGCCAAGAGGAGTATGAGTCTGCTCATATGAGAGCTCTGAGAGCAAGCTCAG 1077  
1374 ATGAAAGAGACACAGAGGCGAGATGAGAAAGTCTCTCTCTGAGCATTAACCCAG 1433  
1078 ATGAAAGAGACACAGAGGCGAGATGAGAAAGTCTCTCTCTGAGCATTAACCCAG 1137  
1434 CTGAGTTGGGTGAGCTGAGAGAACTTCTCTCCAGCCCTGCAACTTTAGCTTGG 1493  
1138 CTGAGTTGGGTGAGCTGAGAGAACTTCTCTCCAGCCCTGCAACTTTAGCTTGG 1197  
1494 TTGAGCTGCTTCCAGCCCTTCCAGCCCGAGACACACAAAGTTGGCCCGCAGCTGC 1553  
1198 TTGAGCTGCTTCCAGCCCTTCCAGCCCGAGACACACAAAGTTGGCCCGCAGCTGC 1257  
1554 GCTGACATTTAGGCAAGTGAATCTGTGTGTAAGGGGCGGTGCAACACTCTTAGAC 1613  
1258 GCTGACATTTAGGCAAGTGAATCTGTGTGTAAGGGGCGGTGCAACACTCTTAGAC 1317  
1614 AGGCCACCACTTAGACAGCCCACTCTGACCGGGTTCCTAGGCTCTCTCTAGG 1673  
1318 AGGCCACCACTTAGACAGCCCACTCTGACCGGGTTCCTAGGCTCTCTCTAGG 1377  
1674 TCCCTCGGCGGAGAGTGTGTTGTTGTTGAGCTGTTTTCGATGATATAGT 1733  
1378 TCCCTCGGCGGAGAGTGTGTTGTTGTTGAGCTGTTTTCGATGATATAGT 1437  
1734 AGTAGAAATGAAATCATTTGATGTAAGAGCTGATGATCTCTCTGAGGCGCT 1793  
1438 AGTAGAAATGAAATCATTTGATGTAAGAGCTGATGATCTCTCTGAGGCGCT 1497  
1794 CACCCAGTTCAAGTCAAGGCTTCAACCGGAGAGGCTTCTCTCTGAGGCGCT 1853  
1498 CACCCAGTTCAAGTCAAGGCTTCAACCGGAGAGGCTTCTCTCTGAGGCGCT 1557  
1854 GTTTCGAGGCTGTTGAGCTAGACATTTAGCTCCGCAATGAGCTCCAGGTTTACA 1913  
1558 GTTTCGAGGCTGTTGAGCTAGACATTTAGCTCCGCAATGAGCTCCAGGTTTACA 1617  
1914 GACAAATTCACAGGCGGTGAGGCGAGGCTGCTTTTATATGCTCCCATTT 1973  
1618 GACAAATTCACAGGCGGTGAGGCGAGGCTGCTTTTATATGCTCCCATTT 1677  
1974 CACAGAGATACACAGGAGCTCGAGAGGAGACAGATAGGACAGAGGCGCCAGCTTGTG 2033  
1678 CACAGAGATACACAGGAGCTCGAGAGGAGACAGATAGGACAGAGGCGCCAGCTTGTG 1737  
2034 CCTAGCAAAATTCAGAGTACAGCTTCACTGAAACAGAGCTGCTTACTGTGCTGTT 2093  
1738 CCTAGCAAAATTCAGAGTACAGCTTCACTGAAACAGAGCTGCTTACTGTGCTGTT 1797  
2094 CCTAGCAAAATTCAGAGTACAGCTTCTGAGGCTTCACTGTGCTTCTAGAAACCAAG 2153

```

Db      1798  CCTCAAGCATTTATTAAGACCTACTGCGGTGCTGGGTTCACTGTGTCTAGGAACCAAG 1857
Qy      2154  AGGGTCCCAAGTCCCTGGGCTCTGCCCCGCTGCTGCCCCACACCTTCTGACACAG 2213
Db      1858  AGGGTCCCAAGTCCCTGGGCTCTGCCCCGCTGCTGCCCCACACCTTCTGACACAG 1917
Qy      2214  CGGTGGGAGGCGGGGAGAGACAGCTGGGAGCCAGAACCTGAGCTGGAGGAGTCCGCA 2273
Db      1918  CCGTGGGAGGCGGGGAGAGAGAGCTGGGAGCCAGAACCTGAGCTGGAGGAGTCCGCA 1977
Qy      2274  GAAAAGCTCAGGGCGGGTCTTCTCTGTTGCCCCGGGATTTGGCTATGCTGGTACACA 2333
Db      1978  GAAAAGCTCAGGGCGGGTCTTCTCTGTTGCCCCGGGATTTGGCTATGCTGGTACACA 2037
Qy      2334  TGTACTCAGGAGTGTGGGTTTGAACCATTAACCAAGGCCCTTGTGATCAGCTTTA 2393
Db      2038  TGTACTCAGGAGTGTGGGTTTGAACCATTAACCAAGGCCCTTGTGATCAGCTTTA 2097
Qy      2394  ACAAGTATATTTTATTTATCTCTAAACATATGAAGTTTAGGGCCCTAAGGA 2453
Db      2098  ACAAGTATATTTTATTTATCTCTAAACATATGAAGTTTAGGGCCCTAAGGA 2157
Qy      2454  CCTTAGTATCTTCTATTTGGGCTTTCTGAGGTTGAGAGGGTAACTTCTCCAG 2513
Db      2158  CCTTAGTATCTTCTATTTGGGCTTTCTGAGGTTGAGAGGGTAACTTCTCCAG 2217
Qy      2514  GTCAACAGCAAGTCTGGGTTGGGAGAGCAAGCTAGGCTGGGCTTCACTATACC 2573
Db      2218  GTCAACAGCAAGTCTGGGTTGGGAGAGCAAGCTAGGCTGGGCTTCACTATACC 2277
Qy      2574  ACGATGTCCTCCCTCTGATGCTTGGGCTTGGGCTTCAAGGCTTGGGATATCTT 2633
Db      2278  ACGATGTCCTCCCTCTGATGCTTGGGCTTGGGCTTCAAGGCTTGGGATATCTT 2337
Qy      2634  GTCCCAACCTCTCCCTAGATCACTGTGAGGGTCCCTGATGATTTGTGTACACAT 2693
Db      2338  GTCCCAACCTCTCCCTAGATCACTGTGAGGGTCCCTGATGATTTGTGTACACAT 2397
Qy      2694  GCCCATGATATACAGATACACAGATGTACACAGATGTACACAGATGTACACAG 2753
Db      2398  GCCCATGATATACAGATACACAGATGTACACAGATGTACACAGATGTACACAG 2457
Qy      2754  GCTCGCATACCTGACCTGACCCGAGCTTGGGCTTGGGCTTGGGCTTAAAGC 2813
Db      2458  GCTCGCATACCTGACCTGACCCGAGCTTGGGCTTGGGCTTGGGCTTAAAGC 2517
Qy      2814  AGCAGCTCAACCTGCTGCTGCTTCCCTCCACCACTGCTGAGGCTTCTGAGCAGA 2873
Db      2518  AGCAGCTCAACCTGCTGCTGCTTCCCTCCACCACTGCTGAGGCTTCTGAGCAGA 2577
Qy      2874  CCAAGTACCTTGGCTGACCCGCTGTGCGCTTCAACCAAGGACAGCCCGCCAC 2933
Db      2578  CCAAGTACCTTGGCTGACCCGCTGTGCGCTTCAACCAAGGACAGCCCGCCAC 2637
Qy      2934  ATGATCTCCGCTGATCACTATCAATAAAGTGGTTTCAAAAAAATTTTAAAAA 2993
Db      2638  ATGATCTCCGCTGATCACTATCAATAAAGTGGTTTCAAAAAAATTTTAAAAA 2697
Qy      2994  AA 2995
Db      2698  AA 2699

```

```

RESULT 8
LOCUS   AX572850
DEFINITION Sequence 3 from Patent M002062849.
ACCESSION AX572850
VERSION  AX572850.1 GI:26004938
KEYWORDS
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS Neelan, B., Ketchum, K.A., di Francesco, V. and Beasley, E.M.
TITLE Isolated human ras-like proteins, nucleic acid molecules encoding
JOURNAL them, and uses thereof
Patent: WO 02062849-A 3 15-AUG-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..11221
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 81.1%; Score 2500.4; DB 6; Length 11221;
Best Local Similarity .99.8%; Pred. No. 0;
Matches 2504; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy      469  TTCAATCCTCAGAGGAGATGTTTCAATCTGTGTTTACCTTGATTAACCGGAGTCTTC 528
Db      7681  TCTCTCCTGAGAGGAGATGTTTCAATCTGTGTTTACCTTGATTAACCGGAGTCTTC 7740
Qy      529  GATGAGTCAAGCGCTTCAAGAGAGATCTGAGTCAAGTCTGCTGTAAGAAACAG 588
Db      7741  GATGAGTCAAGCGCTTCAAGAGAGATCTGAGTCAAGTCTGCTGTAAGAAACAG 7800
Qy      589  ACCAAGAGGCGGCGAGCTGCCATGTCATCTGTGCAACAAAGACAGCCGCGAG 648
Db      7801  ACCAAGAGGCGGCGAGCTGCCATGTCATCTGTGCAACAAAGACAGCCGCGAG 7860
Qy      649  CTGTGCGCCGAGAGTCCCAACCAAGAGGCGAGTGTGTGTGGGAGAGAACTTC 708
Db      7861  CTGTGCGCCGAGAGTCCCAACCAAGAGGCGAGTGTGTGTGGGAGAGAACTTC 7920
Qy      709  GCTTACTTCAAGTGTGCGGCAAGAAACAAACGTCGACAGATGTTCTACGTGCTC 768
Db      7921  GCTTACTTCAAGTGTGCGGCAAGAAACAAACGTCGACAGATGTTCTACGTGCTC 7980
Qy      769  TTCAGATGCGCAAGCTGCCACAGAGATGAGCCCGCTTCATGCGAAGATCTCCGTG 828
Db      7981  TTCAGATGCGCAAGCTGCCACAGAGATGAGCCCGCTTCATGCGAAGATCTCCGTG 8040
Qy      829  CAGTACGATGACGCTTTCACACCCAGAGGCTTTCAGATGCGCGGTCAAGAGATGAC 888
Db      8041  CAGTACGATGACGCTTTCACACCCAGAGGCTTTCAGATGCGCGGTCAAGAGATGAC 8100
Qy      889  GCTATGCGATGATCTGCGCTTGCCTGCGCGCGCCGACGTCACAGATGACCTCAAGTAC 948
Db      8101  GCTATGCGATGATCTGCGCTTGCCTGCGCGCGCCGACGTCACAGATGACCTCAAGTAC 8160
Qy      949  ATCAAGGCGCAAGTCTTTCGGAAGGCCAGGCTCGTGAAGAGGACAAATGCAATTCAG 1008
Db      8161  ATCAAGGCGCAAGTCTTTCGGAAGGCCAGGCTCGTGAAGAGGACAAATGCAATTCAG 8220
Qy      1009  TGAAGGAGGATGCTGTGGGCGGGGCTTGGCCAGTGTCTTCAAGGAGTGTGCGCCAGATGC 1068
Db      8221  TGAAGGAGGATGCTGTGGGCGGGGCTTGGCCAGTGTCTTCAAGGAGTGTGCGCCAGATGC 8280
Qy      1069  CCACTGTGCGCATCTCCCAACGAGGCGCGGACGAGTCTTGTTCACAGACTTAAAGCA 1128
Db      8281  CCACTGTGCGCATCTCCCAACGAGGCGCGGACGAGTCTTGTTCACAGACTTAAAGCA 8340
Qy      1129  CCAAGTGAAGGCGCGGCGGCTGCGCTCGCAATTCGTGCTTCTCAAGCTTTC 1188
Db      8341  CCAAGTGAAGGCGCGGCGGCTGCGCTCGCAATTCGTGCTTCTCAAGCTTTC 8400
Qy      1189  CTGAGTCCGCTTGTTCACAGTCTTGTGTGTTTCAATCTCTCTGTGAGAGACACATCT 1248
Db      8401  CTGAGTCCGCTTGTTCACAGTCTTGTGTGTTTCAATCTCTCTGTGAGAGACACATCT 8460
Qy      1249  CTGAGCTCAAGATTAAGGAGAGACTCAAGTAACTTCTCTCTGTGGGTTGGAAG 1308

```

Db 8461 CTGACGCTCAAGATTAGGAGAGACTCAAGTTACACTTCTCTCTCTG6GTTGAG 8520  
 Qy 1309 AAATGTTATGCCAGAGGGGTGAGATTGCTGCTCATATG3AGCTCTCTG3GACAGCC 1368  
 Db 8521 AAATGTTATGCCAGAGGGGTGAGATTGCTGCTCATATG3AGCTCTCTG3GACAGCC 8580  
 Qy 1369 TCAGATATATAAGACACAGAAAGGACATAGAAAGGCTCTCTCTCTG3GACATACA 1428  
 Db 8581 TCAGATATATAAGACACAGAAAGGACATAGAAAGGCTCTCTCTCTCTG3GACATACA 8640  
 Qy 1429 CCAGCTTGTTGGGTG3GAGCTG3GAGAACTTCTCTCCAGCCCTGCAACTCTTACGC 1488  
 Db 8641 CCAGCTTGTTGGGTG3GAGCTG3GAGAACTTCTCTCCAGCCCTGCAACTCTTACGC 8700  
 Qy 1489 TCTGTTCACTGCTGCTGACCTCTCTCCAGCCCTGCAACTCTTACGC 1548  
 Db 8701 TCTGTTCACTGCTGCTGACCTCTCTCCAGCCCTGCAACTCTTACGC 8760  
 Qy 1549 GCTGGCTGACATGAGCGAGTGAATCTGTTGCTGAAGGGGGGCTG3GACACACTCT 1608  
 Db 8761 GCTGGCTGACATGAGCGAGTGAATCTGTTGCTGAAGGGGGGCTG3GACACACTCT 8820  
 Qy 1609 AGACACGCGCCACACTTACACAGCCCACTCTGACCGGCTCTGACGCTCTCTC 1668  
 Db 8821 AGACACGCGCCACACTTACACAGCCCACTCTGACCGGCTCTGACGCTCTCTC 8880  
 Qy 1669 CTAGTCTCTCTGCGGACAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1728  
 Db 8881 CTAGTCTCTCTGCGGACAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8940  
 Qy 1729 ATAGATAGTAAATGAAATCAATTGACTGTAAGCTGTAAGCTCTCTG3GACAG 1788  
 Db 8941 ATAGATAGTAAATGAAATCAATTGACTGTAAGCTGTAAGCTCTCTG3GACAG 9000  
 Qy 1789 GGCCTACCCAGTTCAGATCAAGGCTCTCAAGCGGAGCGCTCTCTCTCTCTCTCA 1848  
 Db 9001 GGCCTACCCAGTTCAGATCAAGGCTCTCAAGCGGAGCGCTCTCTCTCTCTCTCA 9060  
 Qy 1849 ACAGGTTTCCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1908  
 Db 9061 ACAGGTTTCCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 9120  
 Qy 1909 TTAACAGCAATTCACAAAGCGTGGGGTGGGACAGGAGCTGTTTATATGCTCC 1968  
 Db 9121 TTAACAGCAATTCACAAAGCGTGGGGTGGGACAGGAGCTGTTTATATGCTCC 9180  
 Qy 1969 CATTCACAGAGATACCAAGCTCGAGGGGACACAGATAGACACAGGCCCACT 2028  
 Db 9181 CATTCACAGAGATACCAAGCTCGAGGGGACACAGATAGACACAGGCCCACT 9240  
 Qy 2029 TTTGTTCCCTAGCAAAATTCAGGGTACGCTCCACTTGAACCAAGGCTGCTCTACTGTC 2088  
 Db 9241 TTTGTTCCCTAGCAAAATTCAGGGTACGCTCCACTTGAACCAAGGCTGCTCTACTGTC 9300  
 Qy 2089 TCGTTCTCAAGCATTTATTAAGCACTTACTG3GTTCACTGTTCTTACAGAA 2148  
 Db 9301 TCGTTCTCAAGCATTTATTAAGCACTTACTG3GTTCACTGTTCTTACAGAA 9360  
 Qy 2149 CCAAGAGGGTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2208  
 Db 9361 CCAAGAGGGTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9420  
 Qy 2209 CACAGCGTGGGAGCGGGGAGACAGTGTG3GACCCAGAACTGAGCTTGGGAGGAGTC 2268  
 Db 9421 CACAGCGTGGGAGCGGGGAGACAGTGTG3GACCCAGAACTGAGCTTGGGAGGAGTC 9480  
 Qy 2269 CACAGAAAGCTCAGGGCGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2328  
 Db 9481 CACAGAAAGCTCAGGGCGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9540  
 Qy 2329 CACCATGTACTCAGGAGTGTG3GTTTGAACCATTAACCAAGAGCCCTTGTCTATCAGC 2388  
 Db 9541 CACCATGTACTCAGGAGTGTG3GTTTGAACCATTAACCAAGAGCCCTTGTCTATCAGC 9600

Qy 2389 TCTTAAACAGATATTTTGTATTTTAACTCTCTTAAACATATTGAAGTTTAAAGCCCTTA 2448  
 Db 9601 TCTTAAACAGATATTTTGTATTTTAACTCTCTTAAACATATTGAAGTTTAAAGCCCTTA 9660  
 Qy 2449 AGAATCTTATGATCTCTTATTTGAGTCTTCTTGAAGTTTGAAGAGTTAATCTTCC 2508  
 Db 9661 AGAATCTTATGATCTCTTATTTGAGTCTTCTTGAAGTTTGAAGAGTTAATCTTCC 9720  
 Qy 2509 TCCAGTTCAACAGCAAGTCTGAGGGTGGGAGAAAGCAAGTATGAGGCTTGGCATTCAGTAC 2568  
 Db 9721 TCCAGTTCAACAGCAAGTCTGAGGGTGGGAGAAAGCAAGTATGAGGCTTGGCATTCAGTAC 9780  
 Qy 2569 ATACCAAGATGTCTCTCTCTTATGATGTTGAGGCTTGGGAGCTTGGGAGCTTGGGAG 2628  
 Db 9781 ATACCAAGATGTCTCTCTCTTATGATGTTGAGGCTTGGGAGCTTGGGAGCTTGGGAG 9840  
 Qy 2629 ATCTGTCTTCAACCTCTCTCTTATGATGTTGAGGCTTGGGAGCTTGGGAGCTTGGGAG 2688  
 Db 9841 ATCTGTCTTCAACCTCTCTCTTATGATGTTGAGGCTTGGGAGCTTGGGAGCTTGGGAG 9900  
 Qy 2689 ACCATGCCATGTATATACAGTACACAGATGTACACAGATGTACACAGATGTACACAGATGTAC 2748  
 Db 9901 ACCATGCCATGTATATACAGTACACAGATGTACACAGATGTACACAGATGTACACAGATGTAC 9960  
 Qy 2749 CCCAGCTCTGATACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 2808  
 Db 9961 CCCAGCTCTGATACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 10020  
 Qy 2809 AAAGCAGAGCTCCACCT 2868  
 Db 10021 AAAGCAGAGCTCCACCT 10080  
 Qy 2869 GCAACCAAGTATCTTGTGCTGACACCGGTTGTCGCGCTGACCCAGAGCAGCCCG 2928  
 Db 10081 GCAACCAAGTATCTTGTGCTGACACCGGTTGTCGCGCTGACCCAGAGCAGCCCG 10140  
 Qy 2929 CCACCATGATCTCGTGTACATATCAATTAAGTGGTTGTACAA 2978  
 Db 10141 CCACCATGATCTCGTGTACATATCAATTAAGTGGTTGTACAA 10190

RESULT 9  
 H5569D19  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk  
 On Dec 2, 1998 this sequence version replaced gi:3445463.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep

Genome Center

Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>  
RP4-569D19 is from the library RP4-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

#### VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP4-569D19 it may be shorter because we sequence overlapping sections only once, except for a short overlap.  
The true left end of clone RP4-569D19 is at 1 in this sequence. The true right end of clone RP5-824119 is at 31286 in this sequence.

#### FEATURES

##### Source

1..114771  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/map="q13.1"  
/clone="RP4-569D19"  
/clone\_id="RPC1-4"  
142..175  
/note="MIR3 repeat: matches 162..195 of consensus"  
repeat\_region  
277..290  
/note="2.3 copies 6 mer GGAGAA 28% conserved"  
repeat\_region  
418..427  
/note="3.3 copies 3 mer CTC 20% conserved"  
repeat\_region  
514..524  
/note="2.2 copies 5 mer CAGCA 22% conserved"  
repeat\_region  
756..765  
/note="2.5 copies 4 mer CCCA 20% conserved"  
repeat\_region  
1012..1253  
/note="MIR repeat: matches 50..258 of consensus"  
repeat\_region  
1012..1253  
/note="AluSc repeat: matches 61..304 of consensus"  
repeat\_region  
complement(1466..2088)  
/note="L2 repeat: matches 2487..3183 of consensus"  
repeat\_region  
2302..2611  
/note="AluSg repeat: matches 1..309 of consensus"  
repeat\_region  
2616..2630  
/note="3.8 copies 4 mer AAC 30% conserved"  
repeat\_region  
2691..2706  
/note="2.7 copies 6 mer GTGGCA 25% conserved"  
repeat\_region  
2749..3041  
/note="L2 repeat: matches 3033..3313 of consensus"  
repeat\_region  
3546..3871  
/note="AluX repeat: matches 1..312 of consensus"  
repeat\_region  
3886..4117  
/note="MIR repeat: matches 13..237 of consensus"  
repeat\_region  
complement(4206..4486)  
/note="AluB repeat: matches 1..282 of consensus"  
repeat\_region  
complement(4747..4858)  
/note="MIR repeat: matches 47..163 of consensus"  
repeat\_region  
4882..4959  
/note="L2 repeat: matches 3187..3266 of consensus"  
repeat\_region  
5079..5101  
/note="1.9 copies 12 mer AATTCAATCATG 46% conserved"  
repeat\_region  
5080..5103  
/note="6.0 copies 4 mer ATTC 30% conserved"

misc\_feature  
5108..5568  
/note="match: GSS: Em:A2520278"  
repeat\_region  
5219..5228  
/note="2.8 copies 4 mer TCTT 22% conserved"  
repeat\_region  
5324..5334  
/note="2.2 copies 5 mer CTCTC 22% conserved"  
repeat\_region  
5493..5545  
/note="2.1 copies 25 mer GAAGGAGAGCTTTATTCACGCA 88% conserved"  
repeat\_region  
5693..5990  
/note="AluY repeat: matches 1..300 of consensus"  
repeat\_region  
5999..6140  
/note="AluSg/x repeat: matches 1..136 of consensus"  
repeat\_region  
6147..6202  
/note="MIR repeat: matches 151..205 of consensus"  
repeat\_region  
complement(6238..6292)  
/note="MIR3 repeat: matches 147..203 of consensus"  
repeat\_region  
6330..6339  
/note="10.0 copies 1 mer T 20% conserved"  
repeat\_region  
6342..6352  
/note="2.2 copies 5 mer GTGGC 22% conserved"  
repeat\_region  
6453..7182  
/note="L1M45 repeat: matches 4129..4859 of consensus"  
repeat\_region  
complement(7294..7605)  
/note="AluX repeat: matches 1..312 of consensus"  
repeat\_region  
7630..7640  
/note="2.8 copies 4 mer ATTA 22% conserved"  
repeat\_region  
7786..8087  
/note="AluSg repeat: matches 6..307 of consensus"  
repeat\_region  
8748..8909  
/note="AluYa5a2 repeat: matches 149..310 of consensus"  
repeat\_region  
8910..8965  
/note="28.0 copies 2 mer AT 112% conserved"  
repeat\_region  
9217..9231  
/note="3.0 copies 5 mer GCTTT 21% conserved"  
repeat\_region  
complement(9458..9902)  
/note="match: GSS: Em:AQ222044"  
complement(9546..9904)  
/note="match: GSS: Em:B57571"  
10091..10506  
/note="L1M45 repeat: matches 4845..5273 of consensus"  
repeat\_region  
10508..10592  
/note="AluI/FLAM repeat: matches 1..88 of consensus"  
repeat\_region  
10593..11608  
/note="L1M45 repeat: matches 5265..6298 of consensus"  
repeat\_region  
11640..11653  
/note="14.0 copies 1 mer A 28% conserved"  
repeat\_region  
complement(11931..12080)  
/note="L1M61 repeat: matches 6010..6161 of consensus"  
repeat\_region  
12092..12108  
/note="2.4 copies 7 mer TGTTC 25% conserved"  
repeat\_region  
12150..12164  
/note="7.5 copies 2 mer AG 23% conserved"  
repeat\_region  
complement(12342..12784)  
/note="L1M61 repeat: matches 5573..6011 of consensus"  
repeat\_region  
complement(13431..13736)  
/note="AluX repeat: matches 1..306 of consensus"  
repeat\_region  
13794..14127  
/note="L1R9 repeat: matches 1..270 of consensus"  
repeat\_region  
14128..14415  
/note="AluX repeat: matches 1..289 of consensus"  
repeat\_region  
14416..14737  
/note="L1R9 repeat: matches 270..612 of consensus"  
repeat\_region  
14741..14756  
/note="2.3 copies 7 mer ACCTGCT 23% conserved"  
repeat\_region  
complement(14909..15033)  
/note="FLAM C repeat: matches 2..126 of consensus"  
repeat\_region  
15054..15074  
/note="10.5 copies 2 mer AC 33% conserved"  
repeat\_region  
15083..15640  
/note="L2 repeat: matches 2715..3306 of consensus"  
repeat\_region  
complement(15622..15787)  
/note="MIR repeat: matches 85..256 of consensus"

```
repeat_region 15851..15860
/note="2.5 copies 4 mer AATA 20% conserved"
repeat_region 16065..16079
/note="2.5 copies 6 mer AGCCCC 30% conserved"
repeat_region 16383..16392
/note="2.5 copies 4 mer TTTA 20% conserved"
repeat_region 16487..16510
/note="12.0 copies 2 mer CT 48% conserved"
repeat_region 16510..16526
/note="8.5 copies 2 mer TG 34% conserved"
repeat_region 16553..16686
/note="WLTRL repeat: matches 328..470 of consensus"
misc_feature 16638..16890
/note="match: GSS: Em:B13882"
repeat_region 16850..16976
/note="MIR repeat: matches 16..148 of consensus"
repeat_region 17399..17409
/note="2.2 copies 5 mer TAAAT 22% conserved"
repeat_region 17416..17430
/note="15.0 copies 1 mer A 21% conserved"
repeat_region 17734..17806
/note="LTR16A1 repeat: matches 12..84 of consensus"
repeat_region 17842..18145
/note="LTR16A1 repeat: matches 147..456 of consensus"
repeat_region 18411..18441

Query Match 81.1%; Score 2500.4; DB 9; Length 114771;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2504; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 469 TCCATCCCTCAGAGGAGATCTTCATCCGTGGTTCAGCTGGATPAACGGAGTCCCTC 528
Db 57913 TCTCTCCCTGACGGGAGATGTTCTCATCTGTGTCTAGCTGATACCGAGATCTTC 57912
QY 529 GATGAGTCAAGCGCTTCAGAGCAGATCTGAGGTCAAGTCTCTGCTGTAAGAACAAG 588
Db 57973 GATGAGTCAAGCGCTTCAGAGCAGATCTGAGGTCAAGTCTCTGCTGTAAGAACAAG 58032
QY 589 ACCAAGAGAGCGGCGAGCTGCCATGTCTGTCTGTGSCAACAGAACACACACCGCGAG 648
Db 58033 ACCAAGAGAGCGGCGAGCTGCCATGTCTGTCTGTGSCAACAGAACACACCGCGAG 58092
QY 649 CTGTCCCGCAGGTGCCACACCGAGCGGAGCTGTGTGTGCGGACGACGAGAACTCC 708
Db 58093 CTGTCCCGCAGGTGCCACACCGAGCGGAGCTGTGTGTGCGGACGACGAGAACTCC 58152
QY 709 GCTTACTTCGAGGTGTGCGCCCAAGAAACACCAACGTGAGAGATGTTCTAAGTCTC 768
Db 58153 GCTTACTTCGAGGTGTGCGCCCAAGAAACACCAACGTGAGAGATGTTCTAAGTCTC 58212
QY 769 TTTCAGATGGCCAGCTGCCACACGAGATGAGCCCGCTGTGATGCGAAGATCTCCG 828
Db 58213 TTTCAGATGGCCAGCTGCCACACGAGATGAGCCCGCTGTGATGCGAAGATCTCCG 58272
QY 829 CAGTACGATGAGCGCTTCCACACCGAGCGCTTCTGATGCGCGCTCAAGAGATGAG 888
Db 58273 CAGTACGATGAGCGCTTCCACACCGAGCGCTTCTGATGCGCGCTCAAGAGATGAG 58332
QY 889 GCTTATGAGATGTCTGCGCTTTCGCGCGCGCCAGCGTCAACAGTGACTTCAAGTAC 948
Db 58333 GCTTATGAGATGTCTGCGCTTTCGCGCGCGCCAGCGTCAACAGTGACTTCAAGTAC 58392
QY 949 ATTCAGAGCCAGATCTCTTGGGGAAGGCGAGCCCGTGAAGGGAACAAGTGCACATCC 1008
Db 58393 ATTCAGAGCCAGATCTCTTGGGGAAGGCGAGCCCGTGAAGGGAACAAGTGCACATCC 58452
QY 1009 TGAGCAGAGGATCTGGGCGGCGGCTTGGCAGTGCCTTCAAGAGAGTGGCCCCAGATGC 1068
Db 58453 TGAGCAGAGGATCTGGGCGGCGGCTTGGCAGTGCCTTCAAGAGAGTGGCCCCAGATGC 58512
QY 1069 CCACTGTGCGATCTCCCAACGAGAGCCCGGACGAGATCTTGTTCACAGACTTAAAGCA 1128
Db 58513 CCACTGTGCGATCTCCCAACGAGAGCCCGGACGAGATCTTGTTCACAGACTTAAAGCA 58572
```

```
QY 1129 CCAACTGAGAGGCCCCCGGGCGCTGCGCTCCGACATGTGCTGACCTTCACAGCTTC 1188
Db 58573 CCAACTGAGAGGCCCCCGGGCGCTGCGCTCCGACATGTGCTGACCTTCACAGCTTC 58632
QY 1189 CTGAGTCCGCTTGTCCACAGCTCTTGATGATTCCTCTCTGTGGAGAGACATCT 1248
Db 58633 CTGAGTCCGCTTGTCCACAGCTCTTGATGATTCCTCTCTGTGGAGAGACATCT 58692
QY 1249 CTGAGCTTCAAGATTAAGCAAGACTCAAGTTACCTTCTCTCTGTGGAGAG 1308
Db 58693 CTGAGCTTCAAGATTAAGCAAGACTCAAGTTACCTTCTCTCTGTGGAGAG 58752
QY 1309 AAATGTTGAGCCAGAGGGGTGAAGATTTGCTGCGTCAATATGAGACCTCTGGACAAG 1368
Db 58753 AAATGTTGAGCCAGAGGGGTGAAGATTTGCTGCGTCAATATGAGACCTCTGGACAAG 58812
QY 1369 TCAGATGAAAAAGACACAGAGGCGCAGATGAGAAAGTCTCTCTCTGCGATTAACA 1428
Db 58813 TCAGATGAAAAAGACACAGAGGCGCAGATGAGAAAGTCTCTCTCTGCGATTAACA 58872
QY 1429 CCGAGCTTGTGTTGGTGGGAGCTGGAGAACTTCTCCAGCGCTGCAACTTACGC 1488
Db 58873 CCGAGCTTGTGTTGGTGGGAGCTGGAGAACTTCTCCAGCGCTGCAACTTACGC 58932
QY 1489 TCTGTTTCAAGCTGCTCTGACACCCCTCCACACCCCGACACACACAGATTGGCCCCCA 1548
Db 58933 TCTGTTTCAAGCTGCTCTGACACCCCTCCACACCCCGACACACAGATTGGCCCCCA 58992
QY 1549 GCTGCGCTTGAATTGAGCCAGTGAAGTCTGTGTCTGAAGGGGGGCTGGCCACACTCTC 1608
Db 58993 GCTGCGCTTGAATTGAGCCAGTGAAGTCTGTGTCTGAAGGGGGGCTGGCCACACTCTC 59052
QY 1609 AGACCAAGCCCAACCACTTAAGCAAGCCCACTCTCAAGCGCTTCCAGCTCTCTC 1668
Db 59053 AGACCAAGCCCAACCACTTAAGCAAGCCCACTCTCAAGCGCTTCCAGCTCTCTC 59112
QY 1669 CTAGTCCCTCCGCGCCGACAGTTGTGCTTGTGTGTTGAGCTGTTTGTGTCAATGT 1728
Db 59113 CTAGTCCCTCCGCGCCGACAGTTGTGCTTGTGTGTTGAGCTGTTTGTGTCAATGT 59172
QY 1729 ATAGTAGTAAGATGAATCAATCAATGTAAGTAAAGCTCTAGTACTCCCTCTGGCCAG 1788
Db 59173 ATAGTAGTAAGATGAATCAATCAATGTAAGTAAAGCTCTAGTACTCCCTCTGGCCAG 59232
QY 1789 GCCCTCACCCAGTTCAAGTCCACCGGCTCCACCGGAGCGCTTCTCTGCTCCCA 1848
Db 59233 GCCCTCACCCAGTTCAAGTCCACCGGCTCCACCGGAGCGCTTCTCTGCTCCCA 59292
QY 1849 ACAGGTTTCGTTGCTGTTGAGCTTGAATTTGAGCTTGAATTTGAGCTTGAATTTG 1908
Db 59293 ACAGGTTTCGTTGCTGTTGAGCTTGAATTTGAGCTTGAATTTGAGCTTGAATTTG 59352
QY 1909 TTAACAGAAATTTGACAGCGTGGGTGGAGGCGCAGAGATGCTTTTAAATGCTCC 1968
Db 59353 TTAACAGAAATTTGACAGCGTGGGTGGAGGCGCAGAGATGCTTTTAAATGCTCC 59412
QY 1969 CATTTCAAGAGGATTAACACGAGACTCGAGGGGACACAGATGAGCAACAGGCCCACT 2028
Db 59413 CATTTCAAGAGGATTAACACGAGACTCGAGGGGACACAGATGAGCAACAGGCCCACT 59472
QY 2029 TTGTCCCTTACGAAATTTGAGGTTACAGCTTCACTTAACAGAGTGCCTTACTGTGC 2088
Db 59473 TTGTCCCTTACGAAATTTGAGGTTACAGCTTCACTTAACAGAGTGCCTTACTGTGC 59532
QY 2089 TGTCTTCTCAAGATTTTAAAGCACTTAAGGCTGGGTGCTGTTTCACTGTGCTTAAGAAA 2148
Db 59533 TGTCTTCTCAAGATTTTAAAGCACTTAAGGCTGGGTGCTGTTTCACTGTGCTTAAGAAA 59592
QY 2149 CCAAGAGGATCCAGTCTGTGCTCTGCGCGCTCTGCTCCCAACCACTTCTGACA 2208
Db 59593 CCAAGAGGATCCAGTCTGTGCTCTGCGCGCTCTGCTCCCAACCACTTCTGACA 59652
```

QY 2209 CACAGCGTGGGAGGCGGGAGGAGAGCTGGAGCCAGAACTGAGCTTGGAGGAGATC 2268  
 DB 59653 CACAGCGTGGGAGGCGGGAGGAGAGCTGGAGCCAGAACTGAGCTTGGAGGAGATC 59712  
 QY 2269 CGACAGAAAAGCTCAGGCGGGGCTCTTCTCTTGTGGCCGGGATTGGGCTATGCTGGGATC 2338  
 DB 59713 CGACAGAAAAGCTCAGGCGGGGCTCTTCTCTTGTGGCCGGGATTGGGCTATGCTGGGATC 59772  
 QY 2229 CACCATGACTCAGGAGCTGGGTTTGAACCCATTAACCAAGAGCCCTGTATCATGAC 2388  
 DB 59773 CACCATGACTCAGGAGCTGGGTTTGAACCCATTAACCAAGAGCCCTGTATCATGAC 59832  
 QY 2389 TCTTAACAAGATATTTTGTATTTTAAATCTCTTAACATATTGAAGTTTGAAGCCCTTA 2448  
 DB 59833 TCTTAACAAGATATTTTGTATTTTAAATCTCTTAACATATTGAAGTTTGAAGCCCTTA 59892  
 QY 2449 AGGAACCTTAGTATCTTCTTATTTGGTCTTCTTGAAGTTTGAAGAGGTTAAATTAATTC 2508  
 DB 59893 AGGAACCTTAGTATCTTCTTATTTGGTCTTCTTGAAGTTTGAAGAGGTTAAATTAATTC 59952  
 QY 2509 TCCAGGTTCACACAGAGTCTGTGGTGGGAGAGCAAGCTAGCGCTGGGATTCAGATC 2568  
 DB 59953 TCCAGGTTCACACAGAGTCTGTGGTGGGAGAGCAAGCTAGCGCTGGGATTCAGATC 60012  
 QY 2569 ATACCAAGATGTGCTCCCTCTCTTGTATGCTTGGCCCTGGGGCTTTCAGGGCTTTGGGAC 2628  
 DB 60013 ATACCAAGATGTGCTCCCTCTCTTGTATGCTTGGCCCTGGGGCTTTCAGGGCTTTGGGAC 60072  
 QY 2629 ATCTTGTCTTGAACCTTCTCCCTAGATGATCTGTGAGGGTCCCTGTATGATTTGTGTAC 2688  
 DB 60073 ATCTTGTCTTGAACCTTCTCCCTAGATGATCTGTGAGGGTCCCTGTATGATTTGTGTAC 60132  
 QY 2689 ACCATGCCCATGTAT 2748  
 DB 60133 ACCATGCCCATGTAT 60192  
 QY 2749 CCCAGCTCTGATATACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 2808  
 DB 60193 CCCAGCTCTGATATACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 60252  
 QY 2809 AAAGCAGAGCTTCAACCCCTGCTCTGTCTTCTTCCCTTCCACCACTGACCTGACCTTCTGA 2868  
 DB 60253 AAAGCAGAGCTTCAACCCCTGCTCTGTCTTCTTCCCTTCCACCACTGACCTGACCTTCTGA 60312  
 QY 2869 GCAAGCAGAGTATCTTGGCTGACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2928  
 DB 60313 GCAAGCAGAGTATCTTGGCTGACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60372  
 QY 2929 CCACATGAGATCTCCGTGTACATCTATCAATAAAGTGAGTTTGTATCAAA 2978  
 DB 60373 CCACATGAGATCTCCGTGTACATCTATCAATAAAGTGAGTTTGTATCAAA 60422  
 RESULT 10  
 AC145536 178071 bp DNA linear HTG 19-JUL-2003  
 LOCUS Lemur catla clone LB2-219P16, WORKING DRAFT SEQUENCE, 5 ordered  
 DEFINITION pieces.  
 ACCESSION AC145536  
 VERSION AC145536.1 GI:22996777  
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.  
 SOURCE Lemur catla (ring-tailed lemur)  
 ORGANISM Lemur catla  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Strepsirrhini; Lemnidae; Lemur.  
 REFERENCE 1 (bases 1 to 178071)  
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 178071)  
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Mukherjee, S., Hosseini, R.,  
 Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission  
 JOURNAL Submitted (19-JUL-2003) Genome Sciences, Lawrence Berkeley National  
 Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA  
 COMMENT

Sequence Produced by Berkeley PGA  
 Web site: <http://pga.lbl.gov>  
 Center Code: PGABERK  
 Center Project Name: L135  
 Bac Clone Name: LB2-219P16

This sequence has been compared to sequences of other species  
 using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be  
 viewed at:  
[http://pga.lbl.gov/cgi-bin/search\\_cwgcg4?type=ntvalue&APOL5](http://pga.lbl.gov/cgi-bin/search_cwgcg4?type=ntvalue&APOL5)

The order-orientation of the draft sequence was accomplished by  
 using:  
 Avid (<http://baboon.math.berkeley.edu/avid/>),  
 Lagun (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHBI)

If the Bac Library Name is LB1 to LB4, please see website  
 for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>  
 These libraries are available through the BACPAC Resources Center:  
<http://www.chori.org/bacpac/libraries.htm> as LBML-1 to LBML-4.

#### Summary Statistics:

Sequencing vector: Plasmid; pUC18  
 Chemistry: Dye-terminator Big Dye  
 Assembly program: Phrap version 0.990329.

\* NOTE: This is a 'working draft' sequence.  
 \* consists of 5 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

	1	2	3	4	5
49224: contig of 49224 bp in length	49225	49324: gap of unknown length	49325	65337: contig of 16013 bp in length	65338
65338	65437: gap of unknown length	65438	73997: contig of 8560 bp in length	73998	74097: gap of unknown length
74098	129526: contig of 55429 bp in length	129527	129626: gap of unknown length	129627	178071: contig of 48445 bp in length.

#### FEATURES

source  
 1..178071  
 /organism="Lemur catla"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9447"  
 /clone="LB2-219P16"

#### ORIGIN

Query Match 40.2%; Score 1239.8; DB 2; Length 178071;  
 Best Local Similarity 76.8%; Pred. No. 2e-176;  
 Matches 1925; Conservative 0; Mismatches 442; Indels 140; Gaps 28;

QY 473 TCCACAGAGGAGTCTTCACTCTCTGTTGCTTCACTGATTAACCGGAGTCTTCAATG 532  
 DB 59193 TCCCGGAGGAGGAGTCTTCACTCTGTTGCTTCACTGATTAACCGGAGTCTTCAATG 59252  
 QY 533 AGGTCAAGCGCTTCAAGAGAGATCTTGAAGTCAAGTCTGCTGCTGAAGACAAGACA 592  
 DB 59253 AGGTCAAGCGCTTCAAGAGAGATCTTGAAGTCAAGTCTGCTGCTGAAGACAAGACA 59312  
 QY 593 AGGAGGCGGCGAGCTGCCATGTGTATCTGTGCAACAGAACGACCAAGGAGCTGT 652  
 DB 59313 AGGAGGCGGCGAGCTGCCATGTGTATCTGTGCGCAACAGAACGACCAAGGAGCTGT 59372



QY 653 GCCGCAAGTSCCAACCAAGGCGGAGCTGTGTGTGCGGAGCAAGAACTCGGCT 712  
| | | | |  
Db 59373 GCCCGCAAGTSCCAACCAAGGCGGAGCTGTGTGTGCGGAGCAAGAACTCGGCT 59432  
| | | | |  
QY 713 ACTTCAGAGTGTGCGCCAAAGAAACACCAACCTGTGAGAGAAATTTTCAAGTCTTCA 772  
| | | | |  
Db 59433 ACTTCAGAGTGTGCGCCAAAGAAACACCAACCTGTGAGAGAAATTTTCAAGTCTTCA 59492  
| | | | |  
QY 773 GCATAGCCCAAGCTGTGCAACAGAGATGAGCCCGGCTGATCGAAGATCTCGTGAGT 832  
| | | | |  
Db 59493 GCATAGCCCAAGCTGTGCAACAGAGATGAGCCCGGCTGATCGAAGATCTCGTGAGT 59552  
| | | | |  
QY 833 ACGGTGAGGCTTCCACCCAGGCGCTTGTGATGCGCGCTGCAAGAGATGAGAGCTT 892  
| | | | |  
Db 59553 ACGGTGAGGCTTCCACCCAGGCGCTTGTGATGCGCGCTGCAAGAGAGCTT 59612  
| | | | |  
QY 893 ATGCGATGTCTCGGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 952  
| | | | |  
Db 59613 ACGGATGTGTGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 59672  
| | | | |  
QY 953 AGGCAAGGTCTTTCGCGGAGGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1012  
| | | | |  
Db 59673 AGGCAAGGTCTTTCGCGGAGGCGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCG 59732  
| | | | |  
QY 1013 CGAGGAGTGTGCGGCG 1072  
| | | | |  
Db 59733 CG-GGAGCGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 59791  
| | | | |  
QY 1073 TGTGCGCATCTTCCCAACCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1132  
| | | | |  
Db 59792 TGTGCGCATCTTCCCAACCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 59851  
| | | | |  
QY 1133 ACTGAGGCG 1192  
| | | | |  
Db 59852 C--GGAGGCTTCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 59908  
| | | | |  
QY 1193 GTCCGCTTGTCCACAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1252  
| | | | |  
Db 59909 GAGTCTCTTGTGCA-----TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 59959  
| | | | |  
QY 1253 AGCTCAAGATTTAGGCGAGAGACTCAAGTTACCTTCTCTCTGCGGTTGGAAGAA 1312  
| | | | |  
Db 59960 AGGACAGAGAGACAAAGTTTGG--CAACGCGGAGCTTCTGCGCTTGGGTTGGAAGAA 60017  
| | | | |  
QY 1313 GTTGATGCGAGAGGAGT 1372  
| | | | |  
Db 60018 GTTGATGCGAGAGAGACAAAGGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60077  
| | | | |  
QY 1373 GATGAAAAGACA-CAGAAAGCGAGATGAGAAAGTCTCTCTCTGCGATTAACACC 1431  
| | | | |  
Db 60078 GATGAAAAGACAACCGGAGCGAGACGAGAGAGAGTCTCTCTGTGTGTGTGTGTGT 60134  
| | | | |  
QY 1432 AGCTTGT 1491  
| | | | |  
Db 60135 AGCTTGT 60192  
| | | | |  
QY 1492 GATTCAAGT 1551  
| | | | |  
Db 60193 GATTCAAGT 60245  
| | | | |  
QY 1552 GCGGCTGATGAGCAAGT 1611  
| | | | |  
Db 60246 GCGGCTGATGAGCAAGT 60278  
| | | | |  
QY 1612 CCAAGCCCAACCTTGAACCAAGCCCACTTCTGACCGGCTTCTGAGCTTCTTCTTCT 1671  
| | | | |  
Db 60279 ---GCCGTGCACTTGAACCAAGCCCACTTCTGACCGGCTTCTGAGCTTCTTCTTCT 60334  
| | | | |  
QY 1672 GGTGCTTCTGCGGCG 1731  
| | | | |  
Db 60335 GGTGCTTCTC-----CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60389  
| | | | |  
QY 1732 GTAGTAGAATGAAATCATTTGATCTGTAAAGCTTGTGTGTGTGTGTGTGTGTGTGT 1791  
| | | | |

Db 60390 GCAATGACATAGGAGTCTCTGTATCTGTACAGAGCTTACAGTCAACCTTCTGCGGCG 60446  
| | | | |  
QY 1792 CTCACCAAGTTCAGATTCACGCGCTTCCACCGGAGAGCTTCTCTCTGTCTGTCCCAACA 1851  
| | | | |  
Db 60447 CTCGCCCACTTCAAGTTCACGCGCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 60505  
| | | | |  
QY 1852 GGTGCTTCTGCGGCG 1911  
| | | | |  
Db 60506 GGTGCTTCTGCGGCG 60564  
| | | | |  
QY 1912 CAGACATTCAGACAGAGT 1971  
| | | | |  
Db 60565 CAGACATTCAGACAGAGT 60624  
| | | | |  
QY 1972 TTGACAGAGATTCACACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2031  
| | | | |  
Db 60625 TTGACAGAGATTCACACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60683  
| | | | |  
QY 2032 TCCCTTGAAGAAATTCAGGAGTACAGCTTCACTTGAACAGAGTGTGTGTGTGTGTGT 2091  
| | | | |  
Db 60684 TCCCTTGAAGAAATTCAGGAGTACAGCTTCACTTGAACAGAGTGTGTGTGTGTGTGT 60743  
| | | | |  
QY 2092 TTCTTCAAGCAATTTATTAAGCACTTACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2151  
| | | | |  
Db 60744 TTCTTCAAGCAATTTATTAAGCACTTACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60803  
| | | | |  
QY 2152 TGAAGGTCCTTCACTTGT 2211  
| | | | |  
Db 60804 TGAAGGTCCTTCACTTGT 60850  
| | | | |  
QY 2212 AGCGGTGAGAGGCGGAG 2271  
| | | | |  
Db 60851 AGCGGTGAG 60903  
| | | | |  
QY 2272 CAGAAAAGCTCAGGCGCGGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2331  
| | | | |  
Db 60904 CAG-AAAAGCTCCTGAGGAGGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60962  
| | | | |  
QY 2332 CATGTACTCAGGAGT 2391  
| | | | |  
Db 60963 CATGTACTCAGGAGT 61021  
| | | | |  
QY 2392 TAAAGATATTTTGTATTTTAAATCTCTTAAACATTAATTAAGTTTAAAGGCTTAAAG 2451  
| | | | |  
Db 61022 TAAAGATATTTTGTATTTTAAATCTCTTAAACATTAATTAAGTTTAAAGGCTTAAAG 61080  
| | | | |  
QY 2452 AACCTTGTGTCTTCTTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2511  
| | | | |  
Db 61081 AACCTTGTGTCTTCTTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 61140  
| | | | |  
QY 2512 AGGTCAACAGCAAGT 2571  
| | | | |  
Db 61141 AGGTCAACAGCAAGT 61200  
| | | | |  
QY 2572 CCAGATGTGTCTCTCTGT 2631  
| | | | |  
Db 61201 CCAGATGTGTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 61229  
| | | | |  
QY 2632 TTGTCTTCAACCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2691  
| | | | |  
Db 61230 TTGTCTTCAACCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 61288  
| | | | |  
QY 2692 ATGCCATGTATATTAACAAGTACACAGATGTATACACAGATGTATACATGTCTCAAGCC 2751  
| | | | |  
Db 61289 ATGCCATGTATATTAACAAGTACACAGATGTATACACAGATGTATACATGTCTCAAGCC 61346  
| | | | |  
QY 2752 CAGCTTGTATACCTGTACACCTGTACACCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGT 2811  
| | | | |  
Db 61347 CAGCTTGTATACCTGTACACCTGTACACCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGT 61405  
| | | | |  
QY 2812 GAGAGAGCTTCAACCTTGT 2871  
| | | | |





OY		2680	TTTGGTGATACCATTCATCCCATGTATATATACAAGTACACAGATGTACACACAGATGTACAC	2739
Dd	272730	ACTGTGTACAGGGCGGCCCATGTACTGATCATTGCACACACACAGATGGCAC	-----	272780
OY	2740	ATGCTCCAGCCCCAGCTCTGTCACTATCCGTGCACCCTGCAGCCCAGCCTTGAGCCCTGCTGC		2749
Dd	272781	--CGCCCCCGCCAGCTCTGTGCACACAAACACTGCAGCCAGCCCGCCCGTG--TGTG		272830
OY	2800	TCTGTGCTCAAAGCAGCAGCTTCCAACCTGCTCTGTCTCCCTTTCCCACTCCACTGCTGA		2859
Dd	272837	TCTGTGCTCGAGGAGCAGCTCTGACCTTCGCTTGTCTCCCTCCCTCCCGGAGCCCTCT		272890
OY	2860	GCCCTTCTAGCAGACACAGATACCTTGGCTGCAACCGGTGTGTGGCCCGCTCTCACCCAGGC		2919
Dd	272897	G--AGCAACACAGACAGCAGCACTGGCCACGACAGGTGTGTGCTGC---CCACACAG		272955
OY	2920	ACAGCCCCCGCACCATGTGATCTCCGTGTACACTATCAATAAAGTGGGTTTGTACAA		2978
Dd	272952	CCACAGCCCCCTGCGCTGGGTCTCCGTGTGACCTGTCAATAAATAAGTGGGTTTGTACAA		273010
RESULT 13				
LOCUS	AF134409			
DEFINITION	Rattus norvegicus Rhes protein mRNA, complete cds.	3469 bp	mRNA	linear
ACCESSION	AF134409			
VERSION	AF134409.1	GI:5059121		
KEYWORDS				
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
REFERENCE	1 (bases 1 to 3469) Falk,J.D., Vargiu,P., Foye,P.E., Usui,H., Perez,J., Danielson,P.E., Lerner,D.B., Bernal,J. and Sutcliffe,J.G. Rhes: A Striatum-Specific Ras Homolog Related to Dextral J. Neurosci. Res. (1999) In press			
JOURNAL	Submitted (11-MAR-1999) Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road #MB10, La Jolla, CA 92037, USA			
FEATURES				
source	Location/Qualifiers 1..3469 /organism="Rattus norvegicus" /mol_type="mRNA" /strain="Sprague Dawley CD" /db_xref="taxon:10116" /tissue_type="Brain striatal region" 852..1652 /note="similar to Dextral; Ras-like GTP binding protein; thyroid hormone dependent" /codon_start=1 /product="rhes protein" /protein_id="AAD38928.1" /translation="MMKTLISGNCITLANPARKSYRMVVLGASRVGSKSIIVSLNRF EDQVPTTEDFRKYNIHGMYOQLDIDTSNGHPAPMRRLSILGDVAILVSDLS RESPEVRLOIOLIEVKSCLNKDKTEAKELPMVICGNKDHSILCROVPAMAEELLV SGDENCAFEVSAAKNNTVMEMFYVLFSAKLPHEMSPLHHKITSVOYGDAFPDRPC MKRTVGAAYGVSPARPSPVNSDLTKIKATVLRGQARBRDKSIG"			
CDS				
misc_feature				
polyA_signal				
ORIGIN				
Query Match	26.3%	Score 811.2:	DB 10;	Length 3469;
Best Local Similarity	85.8%;	Pred. No. 3,9e-112;		
Matches 913;	Conservative	0;	Mismatches 148;	Indels 3; Gaps 1;

```

Oy      26  GCCGCGCCGAGCGCGCGCCGAGCGAGCGGAGGATCCCGCGAGTGAACCCGCG  85
      667  GTCCGCTCAGCCCAAGCTCCGGGCGCCGCGCAGAGAGATTTGACCGAGTGAACCTCTGG  726
Oy      86  AGCCACCAAGACTCTGGAGAGCTCGCGGCTGAGAGCAGCAGAGCTCCCGCAGCTCC  145
      727  AGCCCCCGAGCGCCGGAGAGCTCGCGCGCGGAGCGCAGCGCGCTTCCCACTCC  786
Oy      146  CGGCGCTCCAGAGAGCTCTGAGCCGCTGCGAGAGCCCGGCGCCCATTTCCAGCCCC  205
      787  AGGAGCTTCAGAGGAGCCCTCCAGCCGATGCTCGGAGTCCGCTCCGCTTCCAGCCCA  846
Oy      206  GAGCCATGATGAAGCTTTGTCACGCGGAACTGACGCTGATGTCGCCGCAAAAAT  265
      847  GAGCCATATGAGACCTTTGTCACGCGGAACTGACACTCACTGACCTGCGCTAAGAACT  906
Oy      266  CATACCGATGATGATGCTGCGGCTGCTCGGAGTGGGAGAGAGCTCATCTGCTGCT  325
      907  CTTACCGATGATGATGCTGCGGCTGCTCGGAGTGGGAGAGAGCTCATCTGCTGCT  966
Oy      326  TCTTCATGAGCGCTTTGAGAGCAGTACACACCACTGAGAGATTTCCACGTAAG  385
      967  TCTTCATGAGCGCTTTGAGAGCAGTACACACCACTGAGAGATTTCCATTCGCAAG  1026
Oy      386  TATACAAATCGCGCGGAGATGATGACGATGATGATGATGATGATGATGATGATGAT  445
      1027  TGTACAAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  1086
Oy      446  CTTCCCGCGCAGTGGCAGGCTGCTCATCTTCACAGGAGATGCTTCATCTGATGCTCA  505
      1087  CTTCCCGCGCAGTGGCAGGCTGCTCATCTTCACAGGAGATGCTTCATCTGATGCTCA  1146
Oy      506  GCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  565
      1147  GCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  1206
Oy      566  TCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  625
      1207  TCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  1266
Oy      626  GCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  685
      1267  GCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1326
Oy      686  TGTGCTCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  745
      1327  TGTGCTCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1386
Oy      746  TGAAGAGATGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCTGCT  805
      1387  TGAAGAGATGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCTGCTTACGCTGCT  1446
Oy      806  CCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  865
      1447  CCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1506
Oy      866  TGGCGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  925
      1507  TGGCGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1566
Oy      926  GCGTCAAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG  985
      1567  GCGTCAAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG  1626
Oy      986  AGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1042
      1627  AGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1686
Oy      1043  GCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1086
      1687  GCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1730

```

```

RESULT 14
BC036988
LOCUS
DEFINITION
Mus musculus RIKEN cDNA 4930526B11 gene, mRNA (CDNA clone
IMAGE:4988981), partial cds.
ACCESSION
BC036988
VERSION
BC036988.1
KEYWORDS
GI:22477555
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 2824)
REFERENCE
Klausner, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Tooby, Y.K.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, Y.K.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shcherbenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D.,
Bonifard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 2824)
Strauberg, R.
Direct Submission
Submitted (23-AUG-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
REMARK
COMMENT
Contact: MGC help desk
Email: gcapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hghl.nih.gov
Ahter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, D.C.,
McDowell, J., Pearson, R., Stantirip, S., Thomas, P.J., Touchman, J.W.,
Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
Series: IRAX Plate: 81 Row: 9 Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomScan gene prediction.
FEATURES
source
1..2824
/molecule="mRNA"

```

```

/strain="FVB/N"
/db_xref="taxon:10090"
/ciome="IMAGE:4988981"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_id="NCI CGAP_Co24"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
<1..2824
/gene="4930526B1Rik"
/notes="Synonym: Rasd2"
/db_xref="locusid:75141"
/db_xref="MGI:1922391"
<1..956
/gene="4930526B1Rik"
/codon_start=3
/product="4930526B1Rik protein"
/protein_id="AAH36988.1"
/db_xref="GI:2247556"
/db_xref="MGI:1922391"
<1..956
/db_xref="MGI:1922391"
/translation="TRPPYATLLEPPOTPGNSAVGAGSGPHLOELHSPSPRARGS
RCSQPRAMKTLSSGNTLVNPAKNSYRMVVLGASRVKSSIVSRFLNGREDYPT
IEDPHKRVNIHGMVQLDIDTGNHFPAMRLSLITGVFTLVLSRSDDEV
KRLQKOLEVSKCKNKTKEAABLPMVLCGNNDHSELRCVPMMEALVSGDENA
YFEVSARKNTNVMFVYLFPMALPHMSPALHKKISVOYGDAFHPRPFCMRRTKVA
GAYGVSPFARPSVNSDLKIKYIKAKVIREGARBRDKCSIQ"

```

## ORIGIN

Query Match 25.3%; Score 779.8; DB 10; Length 2824;  
 Best Local Similarity 85.8%; Pred. No. 2e-107;  
 Matches 878; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

```

QY 67 CCCCCCGAGTACCCCGGAGCCACACAGACTCTGGAGGCTGGCGGCTGAGACGA 126
DB 12 CCGTACGAGTACCCCGGAGCCCGGAGAGCGCGGAGACTCAGCGCTGCGAGCGCA 71
QY 127 GCGAGCTCCCGGAGCTCCCGGCGCTTCAGGAGCTCTGAGCGCGTGCAGAGGCGCG 186
DB 72 GCGGCTCCCGGAGCTCCAGAGCTTCAGGAGCGCTCCAGCGCTGCGGCTCG 131
QY 187 GCCCGCATTCGCCGCCGAGCCGATGATGAAGATTGTCCAGCGGAGACTGCAGCTC 246
DB 132 TCCCGCTGTTCCAGCCCGAGGCGCATGATGAAGACTTGTCCAGTGGAGACTGCAGCTC 191
QY 247 AGTGTCCCGGCAAAAATCATACCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306
DB 192 AATGTGCTGTGAAGACTCTTACCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 251
QY 307 AGCTCCATGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 366
DB 252 AGCTCCATGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 311
QY 367 GAGGATTCACCGTAAAGTATACCAACATCCGCGGCGCATGTATACAGTGCATCTCTG 426
DB 312 GAGGATTCATTCGAAAGTATACCAACATCCAGCGGCGCATGTATACAGTGCATCTCTG 371
QY 427 GATACCTTGGCAACACCCCTTCCCGCATGCGGAGCTGTGTGTGTGTGTGTGTGTGT 486
DB 372 GACACTCTGGCAACACCCATCTCTGCAATGCGCGGCTCTCTCATCTCTCAAGAGAT 431
QY 487 GTCCTTACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 546
DB 432 GTCCTTACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 491
QY 547 CAGAGCAGATCTGAGAGTCAAGTCTGTGCTGAAGAACAGACCAAGAGGCGGAG 606
DB 492 CAGAAAGATCTGAGAGTCAAGTCTGTGCTGAAGAAATAAACCAAGAGGCGGAG 551
QY 607 CTGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 666
DB 552 CTGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 611
QY 667 ACCACGAGGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 726

```

```

DB 612 GCGATGAGGCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 671
QY 727 GCGAAGAAACACCAACGTTGAGAGATGTTCTTACAGTGTCTTTCAGATGCGCAAGCTG 786
DB 672 GCGAAGAAACACCAACGTTGAGAGATGTTCTTACAGTGTCTTTCAGATGCGCAAGCTG 731
QY 787 CCAACAGAGATGAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 846
DB 732 CCGCATGATGAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 791
QY 847 CACCCAGGCGCTTCTGATGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 906
DB 792 CACCCCGGCGCTTCTGATGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 851
QY 907 CCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 966
DB 852 CCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 911
QY 967 CCGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1023
DB 912 CCGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 971
QY 1024 GCGGCGGCGCTTGTGCGCATGCTTCAAGGAGTGTGCGCGCGCGCGCGCGCGCG 1083
DB 972 GCGAGGCGGCTTGTGCGCATGCTTCAAGGAGTGTGCGCGCGCGCGCGCGCG 1031
QY 1084 CCC 1086
DB 1032 CAC 1034

```

RESULT 15  
 AX393362  
 LOCUS AX393362 3020 bp DNA linear PAT 23-MAR-2002  
 DEFINITION Sequence 292 from Patent WO0210217.  
 ACCESSION AX393362  
 VERSION AX393362.1 GI:19701352  
 KEYWORDS  
 SOURCE  
 ORGANISM Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS St Croix, B., Kinzler, K.W. and Vogelstein, B.  
 TITLE Endothelial cell expression patterns  
 JOURNAL Patent: WO 0210217-A 292 07-FEB-2002;  
 The Johns Hopkins University (US)  
 FEATURES  
 source  
 1..3020  
 /organism="Mus musculus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:10090"

## ORIGIN

Query Match 22.5%; Score 693.2; DB 6; Length 3020;  
 Best Local Similarity 87.0%; Pred. No. 2e-94;  
 Matches 774; Conservative 0; Mismatches 113; Indels 3; Gaps 1;

```

QY 200 AGCCCGAGCATATGAGACTTGTTCACGCGGAACCTGACGCTGAGTGCAGCGCA 259
DB 365 AGCCGAGCATATGAGACTTGTTCACGCGGAACCTGACGCTGAGTGCAGCGCA 424
QY 260 AAAACTCATACCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 319
DB 425 AGAACTCTTACCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 484
QY 320 CTGCTTCTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 379
DB 485 CCGCTTCTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 544
QY 380 GTAAGTATACAATCCGCGGAGCATGTATCAAGCTGTGATCTGTGATCA 439

```

Db 545 GCAAGTGTACACATCCACGGGACATGTACAGCTGATATCTTGAGACCTCTGCA 604  
QY 440 ACCACCCCTTCCCGCCATGCGCAGGCTGTCCATCTTCACAGGGAGTCTTCATCTCTG 499  
Db 605 ACCACCCATTCCTCGCATGCGCGGCTCTCATCTTCACAGAGATGTCTTCATCTCTG 664  
QY 500 TGTTCAGCTGTGATACCGGGAGTCTTCGATGATGATCAAGGCTTCAGAGCAGATCC 559  
Db 665 TGTTCAGCTGTGATAGCGGAGTCTTCGATGATGATCAAGGCTTCAGAGCAGATCC 724  
QY 560 TGAAGTCAAGTCTGCTGAAGAACAGACCAAGAGGCGGAGCTGCCCATGTCA 619  
Db 725 TGAAGTCAAGTCTGCTGAAGATTAACCAAGAGGCGAGAGCTGCCCATGTCA 784  
QY 620 TCTGTGCAACAGAACAGACCAAGGAGCTGTGCGCCAGGTGCCACACCGAGCTG 679  
Db 785 TCTGTGGAACAGAAATGACCAAGTGAAGTGTGCGCCAGGTGCCATGAGGCTG 844  
QY 680 AGCTGCTGTGTGCGGCGACAGAACTCGGCTACTTGAGGTGTGCGCCAGAGAACCA 739  
Db 845 AGCTGCTGTGTGCGGCGACAGAACTCGGCTACTTGAGGTGTGCGCCAGAGAACCA 904  
QY 740 CCAAGTGAACGAGATGTTCTACGTCTCTTCAGCATGAGCCAGCTGCGACAGATGA 799  
Db 905 CTAATGTGAACGAGATGTTCTATGTCTTCAGCATGAGCCAGCTGCGCATGATGA 964  
QY 800 GCCCGCGCTGTGATGCAAGATCTCCGTGCAAGTACGCTGACGCTTCCACCCGAGCCCT 859  
Db 965 GCCCTGCACTGCAACATAGATCTCGTGCAGTACGCGGATGCTTTTCAACCCCGGCT 1024  
QY 860 TCTGATGCGCGCGCTCAAGAGATGAGCGCTATGCGATGCTCGCCCTTCGCGCGC 919  
Db 1025 TCTGATGCGTGCCTAGGCTGCGAGGTGCTATGCAATGCTCAACCTTTGCGCGAC 1084  
QY 920 GCCCAGCGTCAACAGTGACTCAAGTACATCAAGGCGCAAGTCTCTTCGAGAGGCGCAGG 979  
Db 1085 GCCCAGTGTCAACAGTGACTCAAGTACATCAAGGCGCAAGTCTCTTCGAGAGGCGCAGG 1144  
QY 980 CCGGTGAAGAGGACAAAGTGCACATTCAGTGA--GCGAGGATGCTGGGGCGGGCTTG 1036  
Db 1145 CCGGAGAGGAGGACAAAGTGTACATTCAGTGAAGGCGCAGGACATCGGGAGGGGGCTTG 1204  
QY 1037 GCCATGCTTCAGGGAGGTGCGCCCAAGTGCCTGCGCATCTCCC 1086  
Db 1205 GGCAGTGCCTTAAGGAGGTGGCTGAGGATTCCTCACTGCGCACATCCAC 1254

Search completed: June 20, 2004, 06:32:59  
Job time : 11846 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - nucleic search, using frame\_p2n model

Run on: June 20, 2004, 22:13:33 ; Search time 464 Seconds  
(without alignments)  
2626.228 Million cell updates/sec

Title: US-09-778-963A-2

Perfect score: 1377

Sequence: 1 MMKTLSGGNCTLSVPAKNSY.....IKAKVLRGQARERDKCTIQ 266

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3017426 segs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09778963/rnat.18062004.123502.16006/app.query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdt -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODR=LOCAL -OUTFMT=ptc -NORM=ext -HARPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09778963@cgn\_1\_1\_480@rnat.18062004.123502.16006  
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query Match Length DB ID Description  
No. Score

1	1377	100.0	3082	9	US-09-778-963A-1	Sequence 1, Appli
2	1372	99.6	2832	13	US-09-918-715-174	Sequence 174, App
3	1372	99.6	2832	16	US-10-185-035-11	Sequence 4, Appli
4	1372	99.6	2973	13	US-09-918-715-197	Sequence 197, App
5	1304	94.7	3020	13	US-09-918-715-292	Sequence 292, App
6	911	66.2	11221	9	US-09-778-963A-1	Sequence 3, Appli
7	911	66.2	16000	16	US-10-185-035-11	Sequence 11, Appli
8	902	65.5	6996	15	US-10-029-386-2051	Sequence 2051, A
9	879	63.8	1744	15	US-10-197-666A-87	Sequence 87, Appli
10	586	42.6	1021	16	US-10-321-039-13	Sequence 13, Appli
11	569	41.3	523	16	US-10-185-035-12	Sequence 12, Appli
12	511	37.1	753	13	US-10-027-632-25123	Sequence 25123, A
13	511	37.1	753	13	US-10-027-632-25123	Sequence 25123, A
14	511	37.1	753	13	US-10-027-632-25123	Sequence 25123, A
15	511	37.1	753	16	US-10-027-632-25123	Sequence 25123, A
16	511	37.1	753	16	US-10-027-632-25123	Sequence 25123, A
17	511	37.1	753	16	US-10-027-632-25123	Sequence 25123, A
18	464	33.7	368	9	US-09-864-761-21643	Sequence 21643, A
19	373	27.1	405	9	US-09-960-352-10273	Sequence 10273, A
20	344	25.0	1108	15	US-10-037-270-945	Sequence 945, App
21	344	25.0	1108	16	US-10-117-722-945	Sequence 945, App
22	344	25.0	4167	17	US-10-399-455-36	Sequence 36, Appli
23	341	24.8	597	10	US-09-873-546-4	Sequence 4, Appli
24	341	24.8	597	15	US-10-029-386-22936	Sequence 22936, A
25	300	21.8	2223	13	US-10-344-404-22	Sequence 22, Appli
26	297	21.6	1733	13	US-10-231-956A-428	Sequence 428, App
27	297	21.6	1733	15	US-10-096-534-66	Sequence 66, Appli
28	296	21.5	1504	15	US-10-197-666A-81	Sequence 81, Appli
29	296	21.5	2340	15	US-10-106-698-1073	Sequence 1073, Ap
30	296	21.5	3624	15	US-10-084-817-186	Sequence 186, App
31	294.5	21.4	3300	16	US-10-037-270-456	Sequence 456, App
32	294.5	21.4	3300	16	US-10-117-722-456	Sequence 456, App
33	294	21.4	552	15	US-10-210-120-89	Sequence 89, Appli
34	288	20.9	3346	9	US-09-764-868-67	Sequence 67, Appli
35	286.5	20.8	819	16	US-10-369-493-45670	Sequence 45670, A
36	286.5	20.8	930	16	US-09-801-368-285	Sequence 285, App
37	286.5	20.8	930	16	US-10-369-493-46056	Sequence 46056, A
38	286	20.8	604	13	US-10-027-632-106883	Sequence 106883, A
39	286	20.8	604	13	US-10-027-632-106884	Sequence 106884, A
40	286	20.8	604	13	US-10-027-632-106884	Sequence 106885, A
41	286	20.8	604	16	US-10-027-632-106883	Sequence 106883, A
42	286	20.8	604	16	US-10-027-632-106884	Sequence 106884, A
43	286	20.8	604	16	US-10-027-632-106885	Sequence 106885, A
44	286	20.8	612	13	US-10-342-887-1395	Sequence 1395, Ap
45	286	20.8	612	13	US-10-172-118-1395	Sequence 1395, Ap

## ALIGNMENTS

RESULT 1  
US-09-778-963A-1  
Sequence 1, Application US/09778963A  
Patent No. US20020115172A1  
GENERAL INFORMATION:  
APPLICANT: NEBIUM, Beena et al  
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
FILE REFERENCE: CLO01112  
CURRENT APPLICATION NUMBER: US/09/778, 963A  
CURRENT FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 3082  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-778-963A-1

Alignment Scores:  
Pred. No.: 2,056-177  
Score: 1377.00  
Length: 3082  
Matches: 266

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-778-963A-2 (1-266) x US-09-778-963A-1 (1-3082)

QY 1 MetMetlySThrLeuSerSerGlyAsnCythrLeuSerValProAlaIysAsnSerTyr 20  
 DB 211 ATGATGAAGACTTGTCCAGCGGGAACGACAGCTCAGTGTGCCCCGCAAAAACCTCATAC 270  
 QY 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerArgPheLeu 40  
 DB 271 CGCATGGTGTGCTGGTGGCTCTCGGGGGGCAAGCTCCATGCTGTCTGCTTCTC 330  
 QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgIysValTyr 60  
 DB 331 AATGCCCGCTTTGAAGACCAAGTACACCCACCATCGAGACTTCACCGTAAAGTATAC 390  
 QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 DB 391 AACATCCCGCGGACATGTACAGCTCGACATCTTGATACCTCTGGCAACCAACCCCTTC 450  
 QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 DB 451 CCGGCATGCGAGAGCTGTCCATCTCCACAGGGAGTGTCTCATCTGCTGCTACGCTG 510  
 QY 101 AspAsnArgGluSerPheAspGluValIysArgLeuGlnIysGlnIleLeuGluValIys 120  
 DB 511 GATAACCGGAGATCTTCATGATGAGGTCAAGCGCTTCAGAAAGACATCTCGAGGTCAAG 570  
 QY 121 SerCysLeuIysAsnIysThrIysGlnAlaIleGluLeuPheMetValIleCysGlnIysn 140  
 DB 571 TCCTCCCTGAAACAAAGCAAGAGGCGGAGCTGCCCATGCTCATCTGTGTGACAC 630  
 QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuLeuVal 160  
 DB 631 AAGAACGACACGCGGAGCTGTGCCCGGAGTGTGCCACACCGAGGCGAGCTGTGGTG 690  
 QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIysIleAsnThrAsnValAsp 180  
 DB 691 TCGGCGGACGAGAACTCCGCTACTTCAGAGTGTGGCCAGAAAGAACCAACCTGTGAC 750  
 QY 181 GluMetPheTyrValLeuPheSerMetAlaIysLeuProHisGluMetSerProAlaLeu 200  
 DB 751 GAGATGTTCTACGTGCTTTCAGACATGGCCAAAGCTGCCACAGAGTGGCCCCGCTG 810  
 QY 201 HisArgIysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
 DB 811 CATCCCAAGATCTCCGTGCAAGTACGGGTACGCTTCACCCAGGCGCTTCGTGACATGGC 870  
 QY 221 ArgValIysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240  
 DB 871 CGCGTCAAGAGATGGAGCCCTATGAGCATGTCTCGCCCTCGCCGCGCCCAAGCGTC 930  
 QY 241 AsnSerAspLeuIysTyrIleIysAlaIysValLeuArgGluGlyGlnAlaArgGluArg 260  
 DB 931 AACAGTGAACCTCAAGTACATCAAGGCCAAGTCTCTTGGAAGGCCAGGCCGCTGAGAG 990  
 QY 261 AspIysCysThrIleGln 266  
 DB 991 GACAACTGACCATCCAG 1008

## RESULT 2

US-09-918-715-174  
 ; Sequence 174, Application US/09918715  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brad St. Croix  
 ; APPLICANT: Bert Vogelstein  
 ; APPLICANT: Kenneth Kinzler  
 ; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
 ; FILE REFERENCE: 1107.00134

CURRENT APPLICATION NUMBER: US/09/918, 715  
 ;  
 ; CURRENT FILING DATE: 2001-08-01  
 ; PRIOR APPLICATION NUMBER: 60/222,599  
 ; PRIOR FILING DATE: 2000-08-02  
 ; PRIOR APPLICATION NUMBER: 60/224,360  
 ; PRIOR FILING DATE: 2000-08-11  
 ; PRIOR APPLICATION NUMBER: 60/262,850  
 ; PRIOR FILING DATE: 2000-04-11  
 ; NUMBER OF SEQ ID NOS: 358  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 174  
 ; LENGTH: 2832  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-918-715-174

## Alignment Scores:

Pred. No.: 8,65e-177 Length: 2832  
 Score: 1372.00 Matches: 265  
 Percent Similarity: 99.62% Conservative: 1  
 Best Local Similarity: 99.62% Mismatches: 0  
 Query Match: 99.64% Indels: 0  
 DB: 13 Gaps: 0

US-09-778-963A-2 (1-266) x US-09-918-715-174 (1-2832)

QY 1 MetMetlySThrLeuSerSerGlyAsnCythrLeuSerValProAlaIysAsnSerTyr 20  
 DB 37 ATGATGAAGACTTGTCCAGCGGGAACGACAGCTCAGTGTGCCCCGCAAAAACCTCATAC 96  
 QY 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerArgPheLeu 40  
 DB 97 CGCATGGTGTGCTGGTGGCTCTCGGGGGGCAAGCTCCATGCTGTCTGCTTCTC 156  
 QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgIysValTyr 60  
 DB 157 AATGCCCGCTTTGAAGACCAAGTACACCCACCATCGAGACTTCACCGTAAAGTATAC 216  
 QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 DB 217 AACATCCCGCGGACATGTACAGCTCGACATCTTGATACCTCTGGCAACCAACCCCTTC 276  
 QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 DB 277 CCGGCATGCGGAGCTGTCCATCTCCACAGGGAGTGTCTCATCTGCTGCTACGCTG 336  
 QY 101 AspAsnArgGluSerPheAspGluValIysArgLeuGlnIysGlnIleLeuGluValIys 120  
 DB 337 GATAACCGGAGATCTTCATGATGAGGTCAAGCGCTTCAGAAAGCAAGTCTGGAGGTCAAG 396  
 QY 121 SerCysLeuIysAsnIysThrIysGlnAlaIleGluLeuPheMetValIleCysGlnIysn 140  
 DB 397 TCCTCCCTGAAACAAAGCAAGAGGCGGAGTGTGCCCATGCTCATCTGTGTGACAC 456  
 QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuLeuVal 160  
 DB 457 AAGAACGACACGCGGAGCTGTGCCCGGAGTGTGCCACACCGAGGCGAGCTGTGGTG 516  
 QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIysIleAsnThrAsnValAsp 180  
 DB 517 TCGGCGGACGAGAACTCCGCTACTTCAGAGTGTGGCCAAAGATCCCAACCAACCTGTGAC 576  
 QY 181 GluMetPheTyrValLeuPheSerMetAlaIysLeuProHisGluMetSerProAlaLeu 200  
 DB 577 GAGATGTTCTACGTGCTTTCAGATGAGCCAAAGCTGCCACAGATAGAGCCCGCCTG 636  
 QY 201 HisArgIysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
 DB 637 CATGCCAAGATCTCCGTGCAAGTACGGGTACGCTTCACCCAGGCGCTTCGTGACATGGC 696  
 QY 221 ArgValIysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240  
 DB 697 CGCGTCAAGAGATGGAGCCCTATGAGCATGTGTCTGCTTGGCCCCGCGCCCAAGCGTTC 756

```

Db      517 TCGGGGACGAGAACTGGCCCTTACTTCAGAGGTGTGGCCAGAGAAACCAACGTCGAC
Oy      181 GJUmEcPheYrYalLeuPheSeSerMeAlaYsLeuProHISJUmEcSerProAlaLeu
Db      577 GAGATGTCTCAAGTGTCTCTTCAGCANTGGCGAAGCTGCCACACGAGATGAGGCCCGGCTG
Oy      201 HIArGyValLeSeSerValGIntYrGJYAPAlaPheHISProArGProPheCyMeArG
Db      637 CATCGCAAGATCTCCCTGTGACGTACGTCGACGCTTCACCCACAGGCCCTTCGTGATCGC
Oy      221 ArGValLYSUmEcSPAlaTYrGJUmEcValSeRProPheAlaArGArGProSeRVal
Db      697 CGGTCAAGAGATGAGACCGCTATGACANTGTCCTCGCCCTTGCCCGGCCGCCAGGCTC
Oy      241 ASnSeRASPleuYsTYrILeYsAlaYsValLeuArGJUGlYGINAlaArGJUnArG
Db      757 AACAGTGACCTCAAGATCAATCAAGGCCAAGTCTTCGGGAAGGCCAGGCCCGTGAAGG
Oy      261 ASpLYeCYsThrILeGln
Db      817 GACAAGTGACACATCCAG 834

RESULT 4
US-09-918-715-197
; Sequence 197, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 197
; LENGTH: 2973
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-715-197

Alignment Scores:
Pred. NO.:          9,36e-177          Length:          2973
Score:              1372.00           Matches:          265
Percent Similarity: 99.62%           Conservative:      0
Best Local Similarity: 99.62%         Mismatches:       1
Query Match:        99.64%           Indels:           0
DB:                 13                Gaps:             0

US-09-778-963A-2 (1-266) x US-09-918-715-197 (1-2973)
Oy      1 MeTMeTLYsThrLeuSeSerGJYANCYsThrLeuSeRValProAlaYsASnSeTYr
Db      208 ATGATGAAGACTTTGTCCAGCGGGAACCTGCACGCTCACTGTGCCCGCAAAAACCTCATAC
Oy      21  ArGMeRValValLeNGlYAlaSeRrYAlaGJYIYsSeSerILeAlSeRrArGPhLeu
Db      268 CGATGGATGGTGTGGTGGTCTCTCGGGGGGCAAGACTCCATCATGTGTCTCGCTCTC
Oy      41  ASnGJYArGPhGJUnArGIntYrThrProThrILeGJUnArGPhHISArGJYValTYr
Db      328 AATGGCCGCTTTGAGGACCAATACACCAACCAACCAAGCACTTCACCGTAAGATATAC
Oy      61  ASnILeArGJYASpMeTYrGJUnLeuAPILeLeuAPThrSeRrGJYASnHISProPhe

```

```

Db      388 AACATCGGGGAGCATGTACACAGCTCGACATCTCTGGATACCTTGGCAACACCCCTTC 447
QY      81  ProAlaMeCArgArgLeuSer11eleuThrGlyAspValPhe11leuValPheSerLeu 100
Db      448 CCCGCATGCGAGGCTGTCTCATCTCCACAGGGAGTGTCTTCACTCTGGGTTCACGCTTG 507
QY      101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnLysGln11leuGluValLys 120
Db      508 GATACCGGAGATCTCTCGATGAGTCAAGCGCTTCAGAGCAAGATCTCTGGAGGTCAAG 567
QY      121 SerCysLeuLysAsnLysThrLysGluValAlaGluLeuProMetVal11leCysGlyAsn 140
Db      568 TCTGCTCTGAAGAACAAGACCAAGAGGGGGGAGCTGCGCATGTCTGTGTGGCAAC 627
QY      141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160
Db      628 AAGAAAGACCAACGCGACGCTGTCCCGCAGGTGCCACCAACGAGGCCAGGCTGCTGTG 687
QY      161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180
Db      688 TCGGGGAGAGAGACTGCGCTCTCTCGAGGTCTCGGCAAGAGAACCAACGAGTGAG 747
QY      181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
Db      748 GAGATGTTCTACGCTGCTTCTTACGATGCGCAAGCTGCCACAGAGTGAAGCCCGCTG 807
QY      201 HisArgLys11SerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220
Db      808 CATCGCAAGATCTCCGTGAGTACGGTGAAGCGCTTCACCCGAGCGCTTCTGCAATGCC 867
QY      221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgProSerVal 240
Db      868 CCGGTAAAGAGATGAGACCGCTTATGSCATGCTCGCCCTTGGCCGCGCCGACGCTG 927
QY      241 AsnSerAspLeuLysTyr11leuValLysValLeuArgGluGlnAlaArgGluArg 260
Db      928 AACAGTGAACCTCAAGTACATCAAGGCCAAGTCTTCCGGAAGGCGCAGCGCTGAGAAG 987
QY      261 AspLysCysThr11leuGln 266
Db      988 GACAAGTGACCAATCCAG 1005

RESULT 5
US-09-918-715-292
; Sequence 292, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 292
; LENGTH: 3020
; TYPE: DNA
; ORGANISM: Mouse
US-09-918-715-292

Alignment Scores:
Pred. No.: 2,036-167 Length: 3020
Score: 1304.00 Matches: 252
Percent Similarity: 96.24% Conservative: 4
Best Local Similarity: 94.74% Mismatches: 10

```

```

Query Match: 94.70% Indels: 0
DB: 13 Gaps: 0
US-09-778-963a-2 (1-266) x US-09-918-715-292 (1-3020)

QY      1 MetMetLysThrLeuSerSerGlyAsnGlyThrLeuSerValProAlaLysAsnSerTyr 20
Db      376 ATGATGAAGACCTTGTCCAGTGGAACTGCACACTCAATGTGCGCTGTAAGACCTCTAC 435
QY      21 ArgMetVal11leuGluValAsnArgValGlyLysSerSer11leValSerArgPheLeu 40
Db      436 CGCATGAGTGTCTGTGGTGTCTCCGAGTGGCAAGACTCATGTGTCTCCGCTTCTTC 495
QY      41 AsnGlyArgPheGluAspGlnTyrThrProThr11leGluAspPheHisArgLysValTyr 60
Db      436 AATGCGCCCTTGAAGACCAAGTACACGCCCACTATCGAGACTTTCATCGCAAGGTGAC 555
QY      61 Asn11leArgGlyAsnMetTyrGlnLeuAsp11leuAspThrSerGlyAsnHisProPhe 80
Db      556 AACATCCAGGGGACATGTACAGCTGATATCTTGGACACTCTTGGCAACCAACCACTTC 615
QY      81 ProAlaMeCArgArgLeuSer11eleuThrGlyAspValPhe11leuValPheSerLeu 100
Db      616 CCGCATGCGCGCGCTCTCATCTCTCAAGAGATGTCTTCATCTGTGTGTTCAGCTG 675
QY      101 AsnAsnArgGluSerPheAspGluValLysArgLeuGlnLysGln11leuGluValLys 120
Db      676 GATACCGGAGATCTTGTGATGAGGTCAAGGCGCTCCAGAAACAGATCTTGGAGGTCAAG 735
QY      121 SerCysLeuLysAsnLysThrLysGluValAlaGluLeuProMetVal11leCysGlyAsn 140
Db      736 TCTGCTCTGAAGATTAACCAAGAGGACAGAGCTGCCCATGTGATCTGTGGAGAC 795
QY      141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160
Db      796 AAGATGACCAAGATGAGTGTGCGCCAGGCTCTCGCATGAGGCTGAGCTGCTGCTG 855
QY      161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180
Db      856 TCTGTGATGAAGAACTGCGCTTATTTGAGGTGTCAACCAAGAAACACTTAAGTGAAC 915
QY      181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
Db      916 GAGATGTTCTATGCTGTCTTACAGATGCCAAGCTGCCCATGATGAGACCCCTGCACTG 975
QY      201 HisArgLys11SerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220
Db      976 CACCATGAAGATCTCCGTCAGTACGGCGATGCTTTTCAACCCCGGCGCTTCTGATCGT 1035
QY      221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgProSerVal 240
Db      1036 CGCATTAAGTGGCAGAGTGCCTTATGTCATGTCTCACCTTGTGCCAGGCCCGAGTGC 1095
QY      241 AsnSerAspLeuLysTyr11leuValLysValLeuArgGluGlnAlaArgGluArg 260
Db      1096 AACAGTGAACCTCAAGTACATCAAGGCCAAGTCTTACGGAAGGCGCAGAGAGG 1155
QY      261 AspLysCysThr11leuGln 266
Db      1156 GACAAGTGATGACATCCAG 1173

RESULT 6
US-09-778-963a-3
; Sequence 3, Application US/09778963A
; Patent No. US20020115172A1
; GENERAL INFORMATION:
; APPLICANT: NEBLAM, Beena et al
; APPLICANT: NEBLAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: C1001112
; CURRENT APPLICATION NUMBER: US/09/778,963A
; CURRENT FILING DATE: 2001-02-08

```

NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: PaedSeq for Windows Version 4.0  
 SEQ ID NO: 3  
 LENGTH: 11221  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-778-963A-3

## Alignment Scores:

Pred. No.:	1,34e-112	Length:	11221
Score:	911.00	Matches:	176
Percent Similarity:	98.33%	Conservative:	1
Best Local Similarity:	97.78%	Mismatches:	3
Query Match:	9	Indels:	0
DB:	96.16%	Gaps:	0

US-09-778-963A-2 (1-266) x US-09-778-963A-3 (1-11221)

```

QY 87 SerIleuThrGlyAspValPheIleuValPheSerLeuAspAsnArgGluSerPhe 106
    |||
DB 7681 TCTCTCCCTGCAGGGGATGCTTCTCATCTCGGTTCAGGCTGATTAACCGGAGTCTTCC 7740
QY 107 AspGluValLysArgLeuGlnLysGlnIleuGluValLysSerCysLeuLysAsnLys 126
    |||
DB 7741 GATGAGGTCAAGCGCTTCAGAGAGAGATCTCGAGGTCAAGTCTGCTGAAGAACAG 7800
QY 127 ThrLysGluAlaAlaGluLeuProMetValIleCysGlyAsnLysAsnAspHisGlyGlu 146
    |||
DB 7801 ACCAAGAGGCGCGGAGCTGCCATGTGCATCTGTGGCAGAACAGACACGCGCGAG 7860
QY 147 LeuCysArgGlnValProThrThrGluAlaGluLeuValSerGlyAspGluAsnSer 166
    |||
DB 7861 CTGTCCCGCCAGAGTCCCAACACCAAGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 7920
QY 167 AlaTyrPheGluValSerAlaLysLysAsnThrAsnValAspGluMetPheTyrValIleu 186
    |||
DB 7921 GCTTACTTCGAGGTGTGGCGCAAGAAACCAACGTGACGAGATGTTCTACGTGCTC 7980
QY 187 PheSerMetAlaLysLeuProHisGluMetSerProAlaLeuHisArgGlyIleSerVal 206
    |||
DB 7981 TTCACAGATGGCCAGACCTGCCACACAGAGATGAGCCCGCCCTGCATCGCAAGATCTCCG 8040
QY 207 GlnTyrGlyAspAlaPheHisProArgProPheCysMetArgArgValLysGluMetAsp 226
    |||
DB 8041 CAGTACGGTGAAGCTCTTCCAGCCCGGCTTGTGATGCGCGGCTCAAGAGATGAC 8100
QY 227 AlaTyrGlyMetValSerProPheAlaArgArgProSerValAsnSerAspLeuTyr 246
    |||
DB 8101 GCTTATGGCATGTGTCTGCGCCCTTGCAGCGCGCCCAAGGTCAACAGTGAACCTCAAGTAC 8160
QY 247 IleLysAlaLysValLeuArgGluGlyGlnAlaArgGluArgAspLysCysThrIleGln 266
    |||
DB 8161 ATCAAGGCGCAAGGTCTTCTCGGAGAGCGCAGCGCTGAGAGAGAGACAAGTGCACATCCAG 8220

```

## RESULT 7

US-10-185-035-11  
 Sequence 11, Application US/10185035  
 Publication No. US20040005706A1  
 GENERAL INFORMATION:  
 APPLICANT: Kenneth P. Monia  
 TITLE OF INVENTION: ANTISENSE MODULATION OF DEXRASI EXPRESSION  
 FILE REFERENCE: RTS-0377  
 CURRENT APPLICATION NUMBER: US/10/185,035  
 NUMBER OF SEQ ID NOS: 135  
 SEQ ID NO 11  
 LENGTH: 16000  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-185-035-11  
 Alignment Scores:

Pred. No.:	2,39e-112	Length:	16000
Score:	911.00	Matches:	176
Percent Similarity:	98.33%	Conservative:	1
Best Local Similarity:	97.78%	Mismatches:	3
Query Match:	9	Indels:	0
DB:	96.16%	Gaps:	0

US-09-778-963A-2 (1-266) x US-10-185-035-11 (1-16000)

```

QY 87 SerIleuThrGlyAspValPheIleuValPheSerLeuAspAsnArgGluSerPhe 106
    |||
DB 11913 TCTCTCCCTGCAGGGGATGCTTCTCATCTCGGTTCAGGCTGATTAACCGGAGTCTTCC 11972
QY 107 AspGluValLysArgLeuGlnLysGlnIleuGluValLysSerCysLeuLysAsnLys 126
    |||
DB 11973 GATGAGGTCAAGCGCTTCAGAGAGAGATCTCGAGGTCAAGTCTGCTGAAGAACAG 12032
QY 127 ThrLysGluAlaAlaGluLeuProMetValIleCysGlyAsnLysAsnAspHisGlyGlu 146
    |||
DB 12033 ACCAAGAGGCGCGGAGCTGCCATGTGCATCTGTGGCAGAACAGACACGCGCGAG 12092
QY 147 LeuCysArgGlnValProThrThrGluAlaGluLeuValSerGlyAspGluAsnSer 166
    |||
DB 12093 CTGTCCCGCCAGAGTCCCAACACCAAGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 12152
QY 167 AlaTyrPheGluValSerAlaLysLysAsnThrAsnValAspGluMetPheTyrValIleu 186
    |||
DB 12153 GCTTACTTCGAGGTGTGGCGCAAGAAACCAACGTGACGAGATGTTCTACGTGCTC 12212
QY 187 PheSerMetAlaLysLeuProHisGluMetSerProAlaLeuHisArgGlyIleSerVal 206
    |||
DB 12213 TTCACAGATGGCCAGACCTGCCACACAGAGATGAGCCCGCCCTGCATCGCAAGATCTCCG 12272
QY 207 GlnTyrGlyAspAlaPheHisProArgProPheCysMetArgArgValLysGluMetAsp 226
    |||
DB 12273 CAGTACGGTGAAGCTCTTCCAGCCCGGCTTGTGATGCGCGGCTCAAGAGATGAC 12332
QY 227 AlaTyrGlyMetValSerProPheAlaArgArgProSerValAsnSerAspLeuTyr 246
    |||
DB 12333 GCTTATGGCATGTGTCTGCGCCCTTGCAGCGCGCCCAAGGTCAACAGTGAACCTCAAGTAC 12392
QY 247 IleLysAlaLysValLeuArgGluGlyGlnAlaArgGluArgAspLysCysThrIleGln 266
    |||
DB 12393 ATCAAGGCGCAAGGTCTTCTCGGAGAGCGCAGCGCTGAGAGAGACAAGTGCACATCCAG 12452

```

## RESULT 8

US-10-029-386-20951  
 Sequence 20951, Application US/10029386  
 Publication No. US20030194704A1  
 GENERAL INFORMATION:  
 APPLICANT: Penn, Sharron G.  
 APPLICANT: Rank, David R.  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
 FILE REFERENCE: A60MICA-X-2  
 CURRENT APPLICATION NUMBER: US/10/029,386  
 NUMBER OF SEQ ID NOS: 34288  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 20951  
 LENGTH: 696  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 OTHER INFORMATION: MAP TO AL022334.1  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.6  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
 OTHER INFORMATION: EST HUMAN HIT: BT754083.1, EVALUE 0.00e+00  
 OTHER INFORMATION: SWISSPROT HIT: P10301, EVALUE 4.00e-10  
 OTHER INFORMATION: NT HIT: g116168170, EVALUE 0.00e+00  
 US-10-029-386-20951

## Alignment Scores:

Pred. No.: 2,51e-113 Length: 696  
 Score: 902.00 Matches: 174  
 Percent Similarity: 99.43% Conservative: 0  
 Best Local Similarity: 99.43% Mismatches: 1  
 Query Match: 65.50% Indels: 0  
 DB: 15 Gaps: 0

US-09-778-963a-2 (1-266) x US-10-029-386-20951 (1-696)

QY 92 AspValPheIleuValPheSerLeuAspAsnArgLysSerPheAspGluValLysArg 111  
 DB 3 GATGTCTCATCTCGGTTCAGCCTGATACCGGAGTCTTCATAGATGCAAGCCG 62  
 QY 112 LeuGluLysGluIleuGluValLysSerCysLeuLysAsnLysThrLysGluVal 131  
 DB 63 CTTCAGAGAGATCTCTGAGGTCAAGTCTGCTGAGAGACCAAGACGAGCGCGG 122  
 QY 132 GluLeuProMetValIleCysGlyAsnLysAsnAspHisGlyGluLeuCysArgGluVal 151  
 DB 123 GACCTGCCATGGTCTCTGTGGCAACAGACGACGCGAGCTGTGCCCGCAGGTG 182  
 QY 152 ProThrThrGluValLeuLeuValLysSerGlyAspGluAsnSerAlaTyrPheGluVal 171  
 DB 183 CCCACCAACGAGCGCGAGTGTGTGGGCGAGAGAACTGGCGCTTACGAGGTG 242  
 QY 172 SerAlaLysLysAsnThrAsnValAspGluMetPheTyrValLeuPheSerMetAlaLys 191  
 DB 243 TCGGCCAAGAGAACCAACCAAGGAGAGATGTTTACCTGCTCTTCAGCATGGCCAA 302  
 QY 192 LeuProHisGluMetSerProAlaLeuHisArgLysIleSerValGlnTyrGlyAspAla 211  
 DB 303 CTGCCACAGAGATGAGCCCGCCCTGTCATCGCAAGTCTCCGTGCAATGAGTGACGCC 362  
 QY 212 PheHisProArgProPheCysMetArgValLysGluMetAspAlaTyrGlyMetVal 231  
 DB 363 TTCACACCGAGCGCTTGTGTCATGCGCGCGTCAGAGATGGAACCTATGCGCATGGT 422  
 QY 232 SerProPheAlaArgArgProSerValAsnSerAspLeuLysTyrIleLysAlaLysVal 251  
 DB 423 TCGCCCTTGTGCGCGCGCCCAAGCGTCACAGTCAAGTCAAGGCGCAGGTC 482  
 QY 252 LeuArgGluGlyGlnAlaArgLysAspLysCysThrIleGln 266  
 DB 483 CTTCGGAGAGCGCGCGCTGTGAGAGGACAGTGCATCCATCCAG 527

## RESULT 9

US-10-197-666a-87  
 / Sequence 87, Application US/10197666A  
 / Publication No. US20030092037A1  
 / GENERAL INFORMATION:  
 / APPLICANT: ASAH KASEI KABUSIKI KAISYA  
 / TITLE OF INVENTION: Elx1 phosphorylation related gene  
 / FILE REFERENCE: PH-1548US  
 / CURRENT APPLICATION NUMBER: US/10197,666A  
 / PRIOR FILING DATE: 2002-11-18  
 / PRIOR APPLICATION NUMBER: JP 2001-218204  
 / PRIOR FILING DATE: 2001-07-18  
 / PRIOR APPLICATION NUMBER: JP 2001-263450  
 / PRIOR FILING DATE: 2001-08-31  
 / PRIOR APPLICATION NUMBER: JP 2002-012176  
 / PRIOR FILING DATE: 2002-01-21  
 / PRIOR APPLICATION NUMBER: US 60/305,884  
 / PRIOR FILING DATE: 2001-07-18  
 / PRIOR APPLICATION NUMBER: US 60/316,304  
 / PRIOR FILING DATE: 2001-09-04  
 / PRIOR APPLICATION NUMBER: US 60/350,027  
 / PRIOR FILING DATE: 2002-01-23  
 / NUMBER OF SEQ ID NOS: 156  
 / SOFTWARE: Patentin Ver. 2.1  
 / SEQ ID NO 87  
 / LENGTH: 1744

## TYPE: DNA

/ ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: CDS

US-10-197-666a-87

## Alignment Scores:

Pred. No.: 1,59e-109 Length: 1744  
 Score: 879.00 Matches: 172  
 Percent Similarity: 77.70% Conservative: 44  
 Best Local Similarity: 61.87% Mismatches: 16  
 Query Match: 63.83% Indels: 4  
 DB: 15 Gaps: 4

US-09-778-963a-2 (1-266) x US-10-197-666a-87 (1-1744)

QY 1 MetMetLysThrLeuSerSerGlyAsnGlyThrLeuSerValProAlaLysAsnSerTyr 20  
 DB 230 ATGATCAAGAGATGTGCCCGAGCGACTCGGAGCTGATATCCCGCCAGAGACTGCTAT 289  
 QY 21 ArgMetValIleuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40  
 DB 290 CCGATGTCATCTCTGCGCTCGTCCAAAGTGGGCAAGAGCGCATGTGTCCGGCTTCTC 349  
 QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60  
 DB 350 ACCGGCGCTTGAAGAGCGCTTACACGCTTACATGAGAGACTTCCACCGAAGTTCTAC 409  
 QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 DB 410 TCCATCCCGCGGAGAGTCTTACAGTGCACATCTCCGACAGTCCGGACACCCGTTCC 469  
 QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 DB 470 CCGCGCATGCGCGCTCTCCATCTTCACAGAGACGTTTTCATCTGTGTTCAGTGTG 529  
 QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnIleLeuGluValLys 120  
 DB 530 GACACCGCGACTCTTGAAGAGTGCAGGGCTCAGAGCGACAGATCTTGACACCCAG 589  
 QY 121 SerCysLeuLysAsnLysThrLysGluValIleLeuLeuProMetValIleCysGlyAsn 140  
 DB 590 TCTTGTCTCAAGAACAAACCAAGAGAACGTCGAGCGCCCTGTGATATGCGGCGAC 649  
 QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluValIleLeuVal 160  
 DB 650 AAGGTGACCGC---GACTTCTACCGCGAGGTGACACGCGAGATCGACACCTGGTG 706  
 QY 161 SerGlyAsp---GluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnVal 179  
 DB 707 GCGCAGACCCCGACCGCTGCGCTTCTTCAGAGATCTGGCAAGAGAACAGAGAGCTTG 766  
 QY 180 AspGluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAla 199  
 DB 767 GACCAAGATGTTCCGGCGCTCTTGCACATGSCAAAGTCCGACAGAGATAGGCCAGAC 826  
 QY 200 LeuHisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMet 219  
 DB 827 CTGACCGCAGAGTCTCGTGCAGTACTGACAGCTGTGCAAGAG-----GCGGTG 880  
 QY 220 ArgArgValLysGluMet-----AspAla 227  
 DB 881 CGGAACAGAGAGCTGCTCGCGCGGCGAGCGCGCGCGCGCGCGAGACCGCGGAGAGCC 940  
 QY 228 TyrGlyMetValSerProPheAlaArgArgProSerValAsnSerAspLeuLysTyrIle 247  
 DB 941 TTGGCATGTCAGGACCTTGTGCGCGCGCGCGCGAGACAGACAGACCTCATATCATC 1000  
 QY 248 LysAlaLysValIleuArgGluGlyGlnAlaArgGluLysArgLysCysThrIle 265  
 DB 1001 CGCGAGAGAGCGCGCGCGCGCGAGCGCGAGCGCGCGAGAGAGCGCTGTCATC 1054

```
RESULT 10
US-10-321-039-13
; Sequence 13, Application US/10321039
; Publication No. US20040014067A1
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Lukowiak, Andrew
; APPLICANT: Jarvis, Nancy
; APPLICANT: Kurensky, David
; TITLE OF INVENTION: Amplification Methods and Compositions
; FILE REFERENCE: FORS-06960
; CURRENT APPLICATION NUMBER: US/10/321,039
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: 09/998,157
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/329,113
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/360,489
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 759
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (561)..(561)
; OTHER INFORMATION: n can be g or t.
US-10-321-039-13

Alignment Scores:
Pred. No.: 1,01e-69 Length: 1021
Score: 586.00 Matches: 113
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 1
Query Match: 42.56% Indels: 0
DB: 16 Gaps: 0

US-09-778-963A-2 (1-266) x US-10-321-039-13 (1-1021)
QY 153 ThrThrguAlaGluLeuLeuValSerGlyAspGluAsnSerAlaTyrPheGluValSer 172
DB 3 ACCACCGAGCGCGAGCTGCTGTCTGCGGCGACGAGAACTGGCTTACTTCAGAGTGTCTG 62
QY 173 AAlaYslySAenThrAsnValAspGluMetPheTyrValLeuPheSerMetAlaYsleu 192
DB 63 GCCAAGAGAAACACCAACGTGACGAGATGTTCTACGTGCTCTTCAGCATGGCCAAAGCTG 122
QY 193 ProHisGluMetSerProAlaLeuHisArgYslyIleSerValGlnTyrGlyAspAlaPhe 212
DB 123 CCACGCGAGATGAGCCCGCCTGTGATCGCAAGATCTCCGTGCACTTACGAGTACGCCCTTC 182
QY 213 HisProAArgProPheCyMetCArgArgValYsGluMetAspAlaTyrGlyMetValSer 232
DB 183 CACCCCAAGCGCTTGTGATGCGCGCGCTCAAGAGATGAGAGCGCTTATGCGATGCTCTCG 242
QY 233 PropheAlaArgArgProSerValAsnSerAspLeuYsTyrIleYsAlaYsValleu 252
DB 243 CCTTCGCGCCCGCGCCCGCAGCGTCAACAGTCACTCAAGTACATCAAGGCCAAAGTCTCT 302
QY 253 ArgGluGlyGlnAlaArgGluArgAspYsCyThrIleGln 266
DB 303 CGGGAAGGCCAGGCGCGGTGAGAGGAGCAAGTGCACATTCAG 344

RESULT 11
US-10-185-035-12
; Sequence 12, Application US/10185035
; Publication No. US20040005706A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF DEXRAS1 EXPRESSION
```

```
; FILE REFERENCE: RTS-0377
; CURRENT APPLICATION NUMBER: US/10/185,035
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 135
; SEQ ID NO 12
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-185-035-12

Alignment Scores:
Pred. No.: 7.3e-68 Length: 523
Score: 569.00 Matches: 111
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.32% Indels: 0
DB: 16 Gaps: 0

US-09-778-963A-2 (1-266) x US-10-185-035-12 (1-523)
QY 1 MetMetLySThrLeuSerSerGlyAsnCySThrLeuSerValProAlaLeuAsnSerTyr 20
DB 189 ATGATGAAGACTTGTGTCCAGCGGGAACAGCAAGCTCACTGTGCGCCGCAAAAATCATAC 248
QY 21 ArgMetValValLeuGlyAlaSerArgValGlyYsSerSerIleValSerArgPheLeu 40
DB 249 CGCATGTGTGTGTGGTGGTGTCTTCGGGTGGGCAAGACTTCATGCTGTCTGCTTC 308
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLeuValTyr 60
DB 309 AATGCCCGCTTTGAGGACCACTACACCAACCACTCGAGACTTCACCGTAAAGTATAC 368
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
DB 369 AACATCCCGCGGCGCATGTACAGCTCGACATCTCGTGAATCTTGGCAACCAACCCCTTC 428
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
DB 429 CCGGCATGCGCAGGCTGTCTCATCTCTCACAGGGAGATGCTTCATCTCTGTGTTACCTTG 488
QY 101 AspAsnArgGluSerPheAspGluValYsArg 111
DB 489 GATTAACCGGAGTCTCTTGTGATGAGGTCAAGGCG 521

RESULT 12
US-10-027-632-25123/c
; Sequence 25123, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25123
; LENGTH: 753
```



```
TYPE: DNA
ORGANISM: Human
US-10-027-632-25123

Alignment Scores:
Pred. No.: 1,19e-59
Score: 511.00
Percent Similarity: 98.08%
Best Local Similarity: 98.08%
Query Match: 37.11%
DB: 13
Matches: 753
Conservative: 102
Mismatch: 0
Indels: 1
Gaps: 0

US-09-778-963a-2 (1-266) x US-10-027-632-25123 (1-753)

QY 163 AspGluAsnSerAlaTyrPheGluValSerAlaLysValAsnThrAsnValAspGluMet 182
DB 751 GACGAGAACTGCCCTTAATT-GAGGTGTGGCCAGAGAAACACCAAGCGAGCGAGATG 693
QY 183 PheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeuHisArg 202
DB 692 TTCACGTGCTCTTCAGCATGGCCCAAGCTGCCACAGAGTAGCCCCCGCTGCATCCG 633
QY 203 LysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArgVal 222
DB 632 AAGATCTCCGTCGACGATGCGGTGACGCCCTTCCACCCAGGCCCTTGCATGCGCGCGTC 573
QY 223 LysGluMetAspAlaTyrGlyMetValSerProPheAlaArgProSerValAsnSer 242
DB 572 AAGGAGATGACCGCTATGGCATGTCTCGCCCTTGGCCGCCCGCCGCGCAACACT 513
QY 243 AspLeuLysTyrIleLysAlaLysValLeuArgGluGlnAlaArgGluArgAspLys 262
DB 512 GACCTCAAGTACATCAAGCCCAAGGCTCTTCCGGAAGGCCAGGCCGCTGAGAGCAAG 453
QY 263 CysThrIleGln 266
DB 452 TGCACCATCCAG 441

RESULT 13
US-10-027-632-25124/c
; Sequence 25124, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25124
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-25124

Alignment Scores:
Pred. No.: 1.19e-59
Length: 753
Matches: 102
Conservative: 0
Mismatch: 0
Indels: 1
Gaps: 0
```

```
Score: 511.00
Percent Similarity: 98.08%
Best Local Similarity: 98.08%
Query Match: 37.11%
DB: 13
Matches: 102
Conservative: 0
Mismatch: 2
Indels: 1
Gaps: 0

US-09-778-963a-2 (1-266) x US-10-027-632-25124 (1-753)

QY 163 AspGluAsnSerAlaTyrPheGluValSerAlaLysValAsnThrAsnValAspGluMet 182
DB 751 GACGAGAACTGCCCTTAATT-GAGGTGTGGCCAGAGAAACACCAAGCGAGCGAGATG 693
QY 183 PheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeuHisArg 202
DB 692 TTCACGTGCTCTTCAGCATGGCCCAAGCTGCCACAGAGTAGCCCCCGCTGCATCCG 633
QY 203 LysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArgVal 222
DB 632 AAGATCTCCGTCGACGATGCGGTGACGCCCTTCCACCCAGGCCCTTGCATGCGCGCGTC 573
QY 223 LysGluMetAspAlaTyrGlyMetValSerProPheAlaArgProSerValAsnSer 242
DB 572 AAGGAGATGACCGCTATGGCATGTCTCGCCCTTGGCCGCCCGCCGCGCAACACT 513
QY 243 AspLeuLysTyrIleLysAlaLysValLeuArgGluGlnAlaArgGluArgAspLys 262
DB 512 GACCTCAAGTACATCAAGCCCAAGGCTCTTCCGGAAGGCCAGGCCGCTGAGAGCAAG 453
QY 263 CysThrIleGln 266
DB 452 TGCACCATCCAG 441

RESULT 14
US-10-027-632-25125/c
; Sequence 25125, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25125
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-25125

Alignment Scores:
Pred. No.: 1.19e-59
Score: 511.00
Percent Similarity: 98.08%
Best Local Similarity: 98.08%
Query Match: 37.11%
DB: 13
Matches: 753
Conservative: 102
Mismatch: 0
Indels: 1
Gaps: 0
```

US-09-778-963A-2 (1-266) x US-10-027-632-25125 (1-753)

Qy 163 AspGluAnSerAlaTyrPheGluValSerAlaValSerAlaThrAsnValAspGluMet 182  
 Db 751 GACGGAACCTGGCTTACTT-GAGGTGTGGCCAAAGAAACCAACGATGAGATG 693  
 Qy 183 PheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeuHisArg 202  
 Db 692 TTCTACGTCCTCTTACGATGCGCAAGCTGCCACACGAGATGAGCCCTTGATGCG 633  
 Qy 203 LysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArgArgVal 222  
 Db 632 AAGATCTCGTGCAAGTACGATGACGCTTCCACCCCAAGCCCTTGATGAGCGCGTC 573  
 Qy 223 LysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerValAsnSer 242  
 Db 572 AAGGAGATGACGCTTACGATGCTGCTCGCCCTTGCCGCCGCCCAAGCTCAACAGT 513  
 Qy 243 AspleuLysTyrIleLysAlaLysValLeuArgGluGlnAlaArgGluArgAspLys 262  
 Db 512 GACCTCAAGTACATCAAGGCCAAGGTCTTCGGGAAGGCCAGGCCCTGAGAGGACAG 453  
 Qy 263 CysThrIleGln 266  
 Db 452 TGCACCATCCAG 441

# RESULT 15

US-10-027-632-25123/c  
 ; Sequence 25123, Application US/10027632  
 ; Publication No. US20030204075A9

## GENERAL INFORMATION:

APPLICANT: Wang, David G.  
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; FILE REFERENCE: 108827.129  
 ; CURRENT APPLICATION NUMBER: US/10/027,632  
 ; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: US 60/193,483  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: US 60/185,218  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/167,363  
 ; PRIOR FILING DATE: 1999-11-23  
 ; PRIOR APPLICATION NUMBER: US 60/156,358  
 ; PRIOR FILING DATE: 1999-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/146,002  
 ; PRIOR FILING DATE: 1999-08-09  
 ; NUMBER OF SEQ ID NOS: 325720  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 25123  
 ; LENGTH: 753  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-027-632-25123

## Alignment Scores:

Pred. No.: 1,196-59 Length: 753  
 Score: 511.00 Matches: 102  
 Percent Similarity: 98.08% Conservative: 0  
 Best Local Similarity: 98.08% Mismatches: 2  
 Query Match: 37.11% Indels: 1  
 DB: 16 Gaps: 0

US-09-778-963A-2 (1-266) x US-10-027-632-25123 (1-753)

Qy 163 AspGluAnSerAlaTyrPheGluValSerAlaValSerAlaThrAsnValAspGluMet 182  
 Db 751 GACGGAACCTGGCTTACTT-GAGGTGTGGCCAAAGAAACCAACGATGAGATG 693

Qy 183 PheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeuHisArg 202  
 Db 692 TTCTACGTCCTCTTACGATGCGCAAGCTGCCACACGAGATGAGCCCTTGATGCG 633  
 Qy 203 LysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArgArgVal 222  
 Db 632 AAGATCTCGTGCAAGTACGATGACGCTTCCACCCCAAGCCCTTGATGAGCGCGTC 573  
 Qy 223 LysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerValAsnSer 242  
 Db 572 AAGGAGATGACGCTTACGATGCTGCTCGCCCTTGCCGCCGCCCAAGCTCAACAGT 513  
 Qy 243 AspleuLysTyrIleLysAlaLysValLeuArgGluGlnAlaArgGluArgAspLys 262  
 Db 512 GACCTCAAGTACATCAAGGCCAAGGTCTTCGGGAAGGCCAGGCCCTGAGAGGACAG 453  
 Qy 263 CysThrIleGln 266  
 Db 452 TGCACCATCCAG 441

Search completed: June 21, 2004, 00:09:52  
 Job time : 478 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 20, 2004, 20:45:38 ; Search time 85 Seconds  
(without alignments)  
1736.670 Million cell updates/sec

Title: US-09-778-963A-2  
Perfect score: 1377  
Sequence: 1 MMTLISGNTLTVPAKNSV.....IKAKVLRBQARRDKCTIQ 266

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Delext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues  
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09778963/runat.18062004.123500.15921/app\_query.fasta\_1.455  
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=p2n.rnt -MINMATCH=0.1 -LOOPCT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US09778963 @CGN 1 1 69 @runat.18062004.123500.15921 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEIOUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEY TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	879	63.8	1841	4	US-09-053-374A-1
2	876.5	63.7	1689	4	US-09-053-374A-4
3	820	59.5	3079	4	US-09-053-374A-6
4	817	59.3	3986	4	US-09-053-374A-3
5	344	25.0	1108	4	US-09-620-312D-945
6	298	21.6	432	4	US-09-621-976-17745
7	294.5	21.4	3300	4	US-09-620-312D-456
8	285	20.7	615	4	US-08-247-946A-5
9	285	20.7	615	5	PCT-US95-06420-5
10	280.5	20.4	551	4	US-09-765-298A-25
11	280.5	20.4	567	4	US-09-223-588-1
12	280.5	20.4	570	4	US-08-884-866A-2

13	280.5	20.4	4480	4	US-09-167-322-12	Sequence 12, Appl
14	279.5	20.3	570	4	US-08-884-866A-11	Sequence 11, Appl
15	279.5	20.3	5775	4	US-08-306-691B-15	Sequence 15, Appl
16	279.5	20.3	5775	4	US-09-023-655-145	Sequence 1145, Ap
17	279.5	20.3	5775	5	PCT-US93-06251-29	Sequence 29, Appl
18	278.5	20.2	567	4	US-09-223-588-2	Sequence 2, Appl1
19	277.5	20.2	480	4	US-08-884-866A-9	Sequence 9, Appl1
20	277.5	20.2	570	4	US-09-765-298A-27	Sequence 27, Appl1
21	277.5	20.2	607	2	US-08-429-964-85	Sequence 85, Appl
22	274	19.9	2436	1	US-08-306-691B-16	Sequence 16, Appl
23	274	19.9	2436	4	US-09-963-137-161	Sequence 161, App
24	273.5	19.9	2436	4	US-09-963-137-165	Sequence 165, App
25	273.5	19.9	574	2	US-08-429-964-83	Sequence 83, Appl
26	269.5	19.6	3026	4	US-09-963-137-136	Sequence 136, App
27	269.5	19.6	3026	4	US-09-963-137-140	Sequence 140, App
28	255.5	18.6	450	4	US-08-884-866A-10	Sequence 10, Appl
29	240	17.4	2894	4	US-09-976-594-842	Sequence 842, App
30	222.5	16.2	1175	2	US-08-773-423-6	Sequence 6, Appl1
31	218	15.8	606	4	US-09-016-434-1082	Sequence 1082, Ap
32	212.5	15.4	563	3	US-09-385-982-426	Sequence 426, Ap
33	212	15.4	2309	3	US-09-317-3	Sequence 3, Appl1
34	212	15.4	2309	4	US-09-454-818-3	Sequence 3, Appl1
35	206	15.0	600	3	US-09-078-317-1	Sequence 1, Appl1
36	206	15.0	600	4	US-09-454-818-1	Sequence 1, Appl1
37	205.5	14.9	730	4	US-09-300-958A-16	Sequence 16, Appl
38	199	14.5	3497	4	US-09-503-505A-2	Sequence 2, Appl1
39	192	13.9	702	3	US-08-942-976-2	Sequence 2, Appl1
40	192	13.9	702	3	US-09-213-397-2	Sequence 2, Appl1
41	192	13.9	702	3	US-09-416-489-2	Sequence 2, Appl1
42	189.5	13.8	1059	4	US-09-391-741A-9	Sequence 9, Appl1
43	189.5	13.8	1062	4	US-09-391-741A-33	Sequence 33, Appl
44	189	13.7	702	4	US-09-016-434-159	Sequence 159, App
45	189	13.7	1084	4	US-09-673-395A-138	Sequence 138, App

#### ALIGNMENTS

RESULT 1  
US-09-053-374A-1  
Sequence 1, Application US/09053374A  
Patent No. 6462177  
GENERAL INFORMATION:  
APPLICANT: YEN, KWANG-MU  
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: US  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,374A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-514  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1841 base pairs  
TYPE: nucleic acid.  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:

NAME/KEY: CDS  
LOCATION: 255..1097  
US-09-053-374A-1

Alignment Scores:  
Pred. No.: 9,87e-111 Length: 1841  
Score: 879.00 Matches: 172  
Percent Similarity: 77.70% Conservative: 44  
Best Local Similarity: 61.87% Mismatches: 46  
Query Match: 63.83% Indels: 16  
DB: 4 Gaps: 4

US-09-778-963A-2 (1-266) x US-09-053-374A-1 (1-1841)

QY 1 MetMetLysThrLeuSerSerGlyAsnCyThrLeuSerValProAlaLysAsnSerTyr 20  
Db 270 ATGATCAAGAGATGTGCGCCGAGCGACTCGAGCTGAGATATCCGCGCAAGATGCTAT 329  
QY 21 ArgMetValIleuLeuGlyValAspArgValGlyLysSerSerIleValSerArgPheLeu 40  
Db 330 CGCATGTGATCTCTGGCTGCTGTCAGAGGTGGCAAGACGCGCATGTGTGCGCTTCTC 389  
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheIleAsnGlySerValTyr 60  
Db 390 ACCGCGCGCTTGAGAGAGCGCTACAGCGCTTACATCGAGACTTCACCGCAAGTTCTAC 449  
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnIleProPhe 80  
Db 450 TCCATCCCGGGAGAGGTCTACAGAGTTCAGATCTCGACATCTCGGCAACCGCGTTC 509  
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
Db 510 CCGCGCATGCGCGGCTCTTCCATCTCTACAGAGAGCTTTTATCTGTGTTCTAGTCTG 569  
QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnIleLeuGluValLys 120  
Db 570 GACACCGCGACTCTTTCAGAGAGGTGACGCGCTCAGGACAGATCTCTGACACCAAG 629  
QY 121 SerCysLeuLysAsnLysThrIleGluAlaIleGluLeuProMetValIleCysGlyAsn 140  
Db 630 TCTTGCCTCAAGAAACAAACAGAGAAAGCTGACGTGCGCTGCTGCTGCTGCTGCTG 689  
QY 141 LysAsnAspHisGlyLysLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160  
Db 690 AAGGCTGACCGC---GACTTCTACCGGACGAGTGACACAGCGGAGATCGACGCTGCTG 746  
QY 161 SerGlyAsp---GluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnVal 179  
Db 747 GCGCAGACCGCCGAGCTGCGCTTCTTCTGAGATCTGCGCAAGAAAGACGAGCTG 806  
QY 180 ArgGluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAla 199  
Db 807 GACCGAGATGTCGCGCGCTCTTCCGCAATGGCAAGCTGCCAGGAGATGAGCCCAAC 866  
QY 200 LeuHisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMet 219  
Db 867 CTGACCGCGAGGTCTCGGTGCAATCTGCGACGTCGTCACAAAG-----CGCGTGT 920  
QY 220 ArgArgValLysGluMet-----AspAla 227  
Db 921 CGGAACAAGAGAGCTGCGGCGCGGCGGCAAGCGGCGCGCGCGGCGGCGGCGGCGGCGG 980  
QY 228 TyrGlyMetValSerProPheAlaArgArgProSerValAsnSerAspLeuLysTyrIle 247  
Db 981 TTGGCATCTGTGGACCTTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1040  
QY 248 LysAlaLysValIleuArgGluGlnIleAlaArgGluArgAspLysCysThrIle 265  
Db 1041 CGCGAAGAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1094

RESULT 2  
US-09-053-374A-4  
Sequence 4, Application US/09053374A

Patent No. 6462177  
GENERAL INFORMATION:  
APPLICANT: YEN, KWANG-MU  
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: US  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,374A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-514  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1689 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 132..971  
US-09-053-374A-4

Alignment Scores:  
Pred. No.: 1.9e-110 Length: 1689  
Score: 876.50 Matches: 171  
Percent Similarity: 77.98% Conservative: 45  
Best Local Similarity: 61.73% Mismatches: 46  
Query Match: 63.65% Indels: 15  
DB: 4 Gaps: 4

US-09-778-963A-2 (1-266) x US-09-053-374A-4 (1-1689)

QY 1 MetMetLysThrLeuSerSerGlyAsnCyThrLeuSerValProAlaLysAsnSerTyr 20  
Db 147 ATGATCAAGAGATGTGCGCCGAGCGACTCGAGCTGAGATATCCGCGCAAGATGCTAC 206  
QY 21 ArgMetValIleuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40  
Db 207 AGATGTGCTATCTCTGCTGCTTCTTCAAGGTGGCAAGCGGCGCATGCTGCTGCTGCT 266  
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheIleAsnGlySerValTyr 60  
Db 267 ACGGCGCTTGAAGAGCTTACACCCCTTACCATTAAGACTTCAACCGAAAGTTTCTAC 326  
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnIleProPhe 80  
Db 327 TCGATCCGCGCGCAAGCTACCACTTGGACATATCTGACACATCTGGCAATACCGTTT 386  
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
Db 387 CCGCGCATGCGCGCGCTCTATCTCTACAGAGACGTTTTCATCTGTGTGCTGAGCTTCA 446  
QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnIleLeuGluValLys 120  
Db 447 GACAAACGCGACTCTTTCGAGAGAGTGCAAAGGCTCAAGACGAGATCTGAGACACCAAG 506  
QY 121 SerCysLeuLysAsnLysThrIleGluAlaIleGluLeuProMetValIleCysGlyAsn 140

```

Db      507 TCCTGTCTCAGAAACAAACCAAGAGATGACGCTGCTGTCATTGGCGTAAC 566
Qy      141 Lysanasphtsigtgluleucyarglnvalprothrthrglnalagluuleuval 160
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      567 AAAAGGACCGG---GACTTACCGCGAAGTGAAGCAGCGGAGATTGAGACTGCTG 623
Qy      161 SerGlyasp---gluanserlalyrphelgluvalseralalyserthrAsnVal 179
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      624 GGCATGACCTTCAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
Qy      180 AspGluMetPheTyrValLeuPheSerMetAlaIysLeuProHisGluMetSerProAla 199
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      684 GACCAAGTGTCCCTGCGCTCTTGGCCATGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGAC 743
Qy      200 LeuHisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMet 219
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      744 TTGACCGCAAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
Qy      220 ArgArgValIysGluMet-----AspAlaTyr 228
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      798 AGGAAACAAAGACTTTCGCGTGGCGGCGGCGGAGGTGGGCGGCAACCGAGATGCTTT 857
Qy      229 GlyMetValSerProPheAlaIArgArgProSerValAsnSerAspLeuLysTyrIleLys 248
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      858 GGCATCTTGGCGGCTTGTGCTGCGAGACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
Qy      249 AlaIysValLeuArgGlnGlyGlnAlaArgGluArgAspLysCysThrIle 265
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      918 GAGAAACCAAGTGTGACGACGAGCTTAAGAGCAAGAGCGCTGTGTCATC 968

RESULT 3
US-09-053-374A-6
; Sequence 6, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-514
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3079 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-053-374A-6

Alignment Scores:
Pred. No.: 3,14e-102 Length: 3079
Score: 820.00 Matches: 170
Percent Similarity: 66.56% Conservative: 45
Best Local Similarity: 52.63% Mismatches: 47
Query Match: 59.55% Indels: 62

```

```

DB:      4 Gaps: 5
US-09-778-963a-2 (1-266) x US-09-053-374A-6 (1-3079)
Qy      1 MetMetLysThrLeuSerSerGlyAsnCysThrLeuSerValProAlaIysAsnSerTyr 20
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      318 ATGATCAAGAGAGATGTCGCCAAGGACTCTGAAGTGAATGATTCGCGCAAGAACTGTAC 377
Qy      21 ArgMetValIleuGlnValIAserArgValGlyLysSerSerIleValSerArgPheLeu 40
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      378 AGGATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
Qy      41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      438 ACGGCGCGCTTCAGAGAGCGCTTACACCCCTTACATTAAGACTTCCACCGAAAGCTTTTAC 497
Qy      61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      498 TCGATCCGCGCGAGAGCTTACCAAGTTGACATCTGACACATCTGCGCAATCTCGCTTT 557
Qy      81 ProAlaMetArgArgLysSerIle----- 88
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      558 CCCGCCATGCGCGCGCTCTCTATCTTCAAGTGAAGTGGGAGCCGACAGGACCGTGGG 617
Qy      88 ----- 88
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      618 GAGGAATCTGCGGAGCGGAGCGGATGGCGGCTGTGTGCTGCGGCTGCTGCTGCT 677
Qy      89 -----LeuThrGlyAspValPhe 94
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      678 GCTCCGTGTGGAGAGCTGCCCTTCACCTTTCACACTGCTCCCTGTGA-GGAGAGTTTTC 736
Qy      95 IleuValIlePheSerLeuAspAsnArgLysSerPheAspGlnValLysArgLeuGlnLys 114
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      737 ATTCTGTGTTCAGCTTGAACAACCGGACTCTTCAAGAGGTGCAAGAGCTTCAAAACG 796
Qy      115 GlnIleuGlnValIysSerCysLeuLysAsnLysThrLysGlnAlaIleGluLeuPro 134
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      797 CAGATCTTACACACCAAGTCTCTGTCTCAAGAACCAAGAGAAATGCGAGCTGCG 856
Qy      135 MetValIleCysGlyAsnLysAsnAspHisGlyGluLeuCysArgGlnValProThrThr 154
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      857 CTGGTCAATTTGGCGGTAAACAAGGGAGCCG---GACTTCAACCGCAAGTGAACAGCGG 913
Qy      155 GlnAlaGluLeuValSerGlyAsp---GluAsnSerAlaTyrPheGlnValSerAla 173
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      914 GAGATTGACGAGCTGTGGCGCATGACCTCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 973
Qy      174 LysIleAsnThrAsnValAspGluMetPheTyrValLeuPheSerMetAlaIysLeuPro 193
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      974 AAGAGAAATGACAGCTGACAGATGTTCCGTGCGCTTGTGCAATGCGCAAGCTGCTCT 1033
Qy      194 HisGluMetSerProAlaIleuHisArgLysIleSerValGlnTyrGlyAspAlaPheHis 213
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1034 AGCGAGATGAGCTCTGACTTGCACCGCAAGGTGTGTGTGCAATCTGTGACGTCTGAC 1093
Qy      214 ProArgProPheCysMetArgArgValIysGluMet----- 225
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1094 AAAAAG-----GCTTGAAGAGAACAGAGCTTCTGTCGCGGACGAGGTGGGGG 1147
Qy      226 -----AspAlaTyrGlyMetValSerProPheAlaArgArgProSerValAsnSer 242
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1148 GACCAAGGAGATGCTTGTGCAATCTTGGCGCGCTTGTGCTGCAAGCTTACGCTGATAGC 1207
Qy      243 AspLeuLysTyrIleLysAlaIysValIleuArgGlnGlyGlnAlaArgGluArgAspLys 262
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1208 GACTCATGTACATCTGTGAGAAACCAAGTGAACGACGCAAGCTAAGCAAGAGACCGC 1267
Qy      263 CysThrIle 265
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1268 TGTGTCTATC 1276
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4

```

US-09-053-374A-3  
Sequence 3, Application US/09053374A  
Patent No. 6462177  
GENERAL INFORMATION:  
APPLICANT: YEN, KWANG-MU  
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: ONE AMGEN CENTER DRIVE  
CITY: THOUSAND OAKS  
STATE: CA  
COUNTRY: US  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,374A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-514  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3986 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-053-374A-3  
  
Alignment Scores:  
Pred. No.: 1,24e-101 Length: 3986  
Score: 817.00 Matches: 172  
Percent Similarity: 61.89% Conservative: 44  
Best Local Similarity: 49.28% Mismatches: 46  
Query Match: 59.33% Indels: 87  
DB: 4 Gaps: 5  
  
US-09-778-963A-2 (1-266) x US-09-053-374A-3 (1-3986)  
QY 1 MetcetylrthrlseuSerSerGlyAsnCysThrlseuSerValProAlaLysAsnSerTyr 20  
Db 789 ATGATCAAGAGATGTGCGCCGAGCGATCGAGCTGAGATATCCGCCAAGAACTGCTAT 848  
QY 21 ArgmetValIalleuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40  
Db 849 CGCATGGTATCTCTCGCTCGTCGCAAGGTGGCAAGACGGCCATGCTGCGGCTTCTC 908  
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60  
Db 909 ACCGCGCTTCGAGAGCGCTTACACGCTTACATCGAGACTTTCACCGCAAGTTCAC 968  
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
Db 969 TCCATCCGCGGAGAGTCTACAGACTCGACATCTCGACAGCTCGGGCAACACCGCTTC 1028  
QY 81 ProAlaMetArgArgLeuSerIleLeuThr----- 90  
Db 1029 CCGCGCATGCGCGCTCTCCATCTCAAGGTGAGCGCGGGGCGGAGGTGCGGGAG 1088  
QY 90 ----- 90  
Db 1089 GGAAGGCGGGGGAACCTTCGCGCAGGGGCGCCGCGAGAGCGGATCGGCTGCTGCGCGC 1148  
QY 90 ----- 90  
Db 1149 CGAGTAGTGGCTTGGCGCTTAGAGAGGTAGCGGCGCGCGCGGCGGCTCAAAAGTCAGCC 1208

QY 90 ----- 90  
Db 1209 CGACTTGCCCTGGGGGCGGACCTCACTTCTCTTCTGCTCTGTGCCCCCTCT 1268  
QY 91 -GlyAspValPheIleLeuValPheSerLeuAspAsnArgLysPheAspGluVally 110  
Db 1269 AGGAGACCTTTTCATCTCGGTGTTCAAGTCTGAGCAACCGGACCTCTTCGAGAGGTGCA 1328  
QY 110 sArgLeuGlnLysGlnIleLeuGluValLysSerCysLeuLysAsnLysThrLysGluAl 130  
Db 1329 GCGGCTCAGCAGCAGATCTCCGACACCACTTCTGCTCAAGAAACAAACAAAGAGGA 1388  
QY 130 aAlaGluLeuProMetValIleCysGlyAsnLysAsnAspHisGlyLeuLysCysArgG 150  
Db 1389 CGTGACGTGCCCCCTGGTATCTCGGCAACAAAGGTGAGCCGCG---GACTTCTACCGCGA 1445  
QY 150 nValProThrThrGlnIleGluLeuValSerGlyAsp---GluAsnSerAlaTyrPhe 169  
Db 1446 GTGGACCAAGCGGAGATCGACAGCTGTGGGAGAGACCCGAGCGCTGCGCTACTT 1505  
QY 169 eGluValSerAlaLysLysAsnThrAsnValAspGluMetPheTyrValLeuPheSerMe 189  
Db 1506 CGAGATCTCGGCAAGAAAGACAGACGCTGAGCCAGATGTTCCGCGGCTCTTCGCGCAT 1565  
QY 189 tAlaLysLeuProHisGluMetSerProAlaLeuHisArgLysIleSerValGlnTyrG 209  
Db 1566 GGCAGAGCTGCCCGAGAGATGAGCCGACCTTCAACGCAAGGTCTCGGTGAGTACTG 1625  
QY 209 yAspAlaPheHisProArgProPheCysMetArgArgValLysGluMet----- 225  
Db 1626 CGAGCTGTGCAACAAGAG-----GGGCTCGGAAACAAGAGGTGCTGCGGCGCGAG 1679  
QY 226 -----AspAlaTyrGlyMetValSerProPheAlaArgAr 237  
Db 1680 CGCGCGCGCGCGCGGAGACCCGCGGAGCGCTTGGCATCTGGACACCTTCGCGCGCG 1739  
QY 237 gProSerValAsnSerAspLeuLysTyrIleLysAlaLysValLeuArgGluGlnAl 257  
Db 1740 GCCAGGCTTACACAGCGACCTCATGTATCATCTCGGAGAGGCCAGCGCGGACGCGAGC 1799  
QY 257 aArgGluArgAspLysCysThrIle 265  
Db 1800 CAAGGACAAGAGGCGCTGCTGCATC 1824  
  
RESULT 5  
US-09-620-312D-945  
Sequence 945, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Weinman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yundong  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tiliugnasht  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FIDE REFERENCE: 784CIP28  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317



```

; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 945
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (172)..(771)
US-09-620-312D-945

Alignment Scores:
Pred. No.: 2.6e-37 Length: 1108
Score: 344.00 Matches: 77
Percent Similarity: 59.47% Conservative: 36
Best Local Similarity: 40.53% Mismatches: 59
Query Match: 24.98% Indels: 18
DB: Gaps: 4

US-09-778-963A-2 (1-266) x US-09-620-312D-945 (1-1108)

QY 18 AsnSerTyrArgMetValIleuGlyAlaSerArgValGlyLysSerSerIleValSer 37
DB 187 AACGATTACCGGCGGCGGCTGTTGGGCTGGCGGCTGGCGAGGCTCCCTGGTGTG 246
QY 38 ArgPheLeuAsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArg 57
DB 247 AGGTTTGTGAAGGACACATTCGCGGAGAGCTACATCCGAGCGTGAAGAACCTACCGG 306
QY 58 LysValTyrAsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsn 77
DB 307 CAATGTGATCGCTGTGACAGAGCATATGCATTGCAGATCCAGACGAGCGGGAGC 366
QY 78 HisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuVal 97
DB 367 CACGAGTCCCGGCGGCGGCTGTGCTTCATCTCCAAAGGCGACGCTTCACTGTG 426
QY 98 PheSerLeuAspAsnArgGluSerPheAspGluValIleuValArgLeuGlnIleLeu 117
DB 427 TACTCCATTACACCGGACGCTCTTGGAGAGGCTCAAGCCCATCTACGAACAATCTGC 486
QY 118 GluValIysSerCysLeuLysAsnLysThrIleGluAlaIleGluLeuPrometValIle 137
DB 487 GAGATCAAGG-----GACGTGAGAGGATCCCATCATGTCTG 525
QY 138 CysGlyAsnLysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGlu 157
DB 526 GTGGGGAACAAGTGTGATGAGAGCCCC--AGCCGCGAGGTGCAAGAGCAGAGCGGAG 582
QY 158 LeuLeuValSerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysAsnThr 177
DB 583 GCCTTG---GCCCGACATGAGAGTGCCTTCATGAGAGCTCAGCCAACTCAACAT 639
QY 178 AsnValAspGluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSer 197
DB 640 AACGTGAAGGAGCTTTTCCAGAGCTGCTCAACCTGAGAG----- 681
QY 198 ProAlaLeuHisArgLysIleSerValGln 207
DB 682 -----CGCAGAGCCGTGATCTCAG 702

RESULT 6
US-09-621-976-17745
; Sequence 17745, Application US/09621976
; Patent No. 6539063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

```

```

; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent .pm
; SEQ ID NO 17745
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17745

Alignment Scores:
Pred. No.: 1.23e-31 Length: 432
Score: 298.00 Matches: 61
Percent Similarity: 64.29% Conservative: 29
Best Local Similarity: 43.57% Mismatches: 42
Query Match: 21.64% Indels: 8
DB: Gaps: 2

US-09-778-963A-2 (1-266) x US-09-621-976-17745 (1-432)

QY 18 AsnSerTyrArgMetValIleuGlyAlaSerArgValGlyLysSerSerIleValSer 37
DB 2 AACGATTACCGGCGGCGGCTGTTGGGCTGGCGGCTGGCGAGGCTCCCTGGTGTG 61
QY 38 ArgPheLeuAsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArg 57
DB 62 AGGTTTGTGAAGGACACATTCGCGGAGAGCTACATCCGAGCGTGAAGAACCTACCGG 121
QY 58 LysValTyrAsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsn 77
DB 122 CAATGTGATCGCTGTGACAGAGCATATGCATTGCAGATCCAGACGAGCGGGAGT 181
QY 78 HisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuVal 97
DB 182 CACGAGTCCCGGCGGCGGCTGTGCTTCATCTCCAAAGGCGACGCTTCACTGTG 241
QY 98 PheSerLeuAspAsnArgGluSerPheAspGluValIleuValArgLeuGlnIleLeu 117
DB 242 TACTCCATTACACCGGACGCTCTTGGAGAGGCTCAAGCCCATCTACGAACAATCTGC 301
QY 118 GluValIysSerCysLeuLysAsnLysThrIleGluAlaIleGluLeuPrometValIle 137
DB 302 GAGATCAAGG-----GACGTGAGAGGATCCCATCATGTCTG 340
QY 138 CysGlyAsnLysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGlu 157
DB 341 GTGGGGAACAAGTGTGATGAGAGCCCC--AGCCGCGAGGTGCAAGAGCAGAGCGGAG 397

RESULT 7
US-09-620-312D-456
; Sequence 456, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aiding J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jiao-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides

```

FILE REFERENCE: 784CIP28  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pc\_fl\_genes version 1.0  
SEQ ID NO 456  
LENGTH: 3300  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (25)..(576)  
US-09-620-312D-456

## Alignment Scores:

Pred. No.:	1,046-29	Length:	3300
Score:	294.50	Matches:	72
Percent Similarity:	56.41%	Conservative:	38
Best Local Similarity:	36.92%	Mismatches:	67
Query Match:	21.39%	Indels:	18
DB:	4	Gaps:	4

US-09-778-963a-2 (1-266) x US-09-620-312D-456 (1-3300)

QY 20 TyrArgMetValIalleuGlyAlaSerArgValGlySerSerIleValSerArgPhe 39  
DB 34 TACAAAGTGTGTCTGTCTGGGCTCGGGCGGGGTAGGCAATCCGCTGACCGTGAAGTTTC 93  
QY 40 LeuAenGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgIleVal 59  
DB 94 GTACCCGGACCTTATTCAGAAATACGACCCACCATCGAGACTTCTACCGCAAGAG 153  
QY 60 TyraenIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 79  
DB 154 ATCGAGTGGATTCGTCCGCGGTGCTGAGATCTTGAGACACCGCGGACCGACGAG 213  
QY 80 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 99  
DB 214 TTCGGTCCATCGGGAGCCTGTACATCAAGACGCGCTCATCTCGCTACACG 273  
QY 100 LeuAspAsnArgGluSerPheAspGlnValIysArgLeuGlnIleLeuGlnVal 119  
DB 274 CTCGTCAACACGACGAGCTTCCAGGACATCAAGCCATCGGGACCAATCATCCGCTG 333  
QY 120 LysSerCysLeuIysAsnIleThrIysGlnIleAlaGluLeuProMetValIleCysGly 139  
DB 334 AAGCGGTATGAGAAA-----GTGCCAGTCACTGTGTTGG 369  
QY 140 AsnIysAsnAspPheIleGlyIleuLeuCysArgGlnValProThrThrGluAlaGluLeu 159  
DB 370 AACAAAGTGACCTG--GAAAGTAGAGAGAGATGCTCAGGAGGAGGAGAGCCCTT 426  
QY 160 ValSerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIysIleAsnThrAsnVal 179  
DB 427 GCTGAA---GAGTGGGGCTGCCCCCTTTATGAAATCTCCGCTAACAGTAAACAAATGTTG 483  
QY 180 AspGluMetPhe-Tyr-----ValLeuPheSerMetAlaIysIle 192  
DB 484 GACGAACCTTTTGCAAAATTTGTGAGGAGCATGAATGATGCTGCTCAGCCGACAAAGAT 543  
QY 192 uProHisGluMetSerProAlaLeuHisArgIleSerVal 206  
DB 544 GACCAATGCTGTTCGATGTAACATACATAGCATCCAAATA 586

## RESULT 8

US-08-247-946A-5  
Sequence 5, Application US/08247946A  
Patent No. 5792638  
GENERAL INFORMATION:

APPLICANT: AARONSON, S.A.; CHAN, A.;  
APPLICANT: MIKI, T.  
TITLE OF INVENTION: NOVEL HUMAN P450-RELATED  
TITLE OF INVENTION: ONCOGENES UNMASKED BY EXPRESSION OF CDNA  
TITLE OF INVENTION: CLONING  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSER: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,946A  
FILING DATE: 24-MAY-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36,434  
REFERENCE/DOCKET NUMBER: 2026-4150  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 615  
TYPE: Nucleic acid  
STRANDEDNESS: Double  
TOPOLOGY: Unknown  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: No  
ORIGINAL SOURCE:  
ORGANISM: Human  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLER:  
FEATURE:  
NAME/KEY: TC21 gene  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-247-946A-5

## Alignment Scores:

Pred. No.:	1,366-29	Length:	615
Score:	285.00	Matches:	74
Percent Similarity:	53.40%	Conservative:	28
Best Local Similarity:	38.74%	Mismatches:	75
Query Match:	20.70%	Indels:	14
DB:	1	Gaps:	4

US-09-778-963a-2 (1-266) x US-08-247-946A-5 (1-615)

QY 17 LysAsnSerTyrArgMetValIalleuGlyAlaSerArgValGlySerSerIleVal 36  
DB 34 CAGAGAAATACCGGCTCGTGTGTCGGGGGGGGGTGGGCAAGTGGGCTCAC 93  
QY 37 SerArgPheLeuAsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHis 56  
DB 94 ATCCAGTTCATCCAGTCTATTGTTGTAACGGATTATGATCCACATGGAATCTTAC 153

```

Qy 57 ArglyValIYrAenIleArgGlyAspMetYrGlnLeuAspIleLeuAspThrSerGly 76
Db 154 ACAAGAGCGTGTGATGATGACAGACAGACCCGGCTAGATATTGATGATACAGACGA 213
Qy 77 AsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeu 96
Db 214 NNNGAAGATTGGAGCATGAGAAACATATATAGAGCATGGCGAAGCTTCCTGTTG 273
Qy 97 ValPheSerLeuAspAsnArgGluSerPheAspGluValAsArgLeuGlnIle 116
Db 274 GTCTTTTCAGTACAGATGAGGCGAGTTTGAAGAACTTAACTTCAAGACAGATT 333
Qy 117 LeuGluValIYsSerCysLeuIYsAsnIYsThrIYsGluAlaIleGluLeuProMetVal 136
Db 334 CTCAGAGTA-----AAGATCGTGAAGATTTCCCAATGATT 369
Qy 137 IleCysGlyAsnIYsAsnAspHisGlyGluLeuCysArgIYsValProThrGluAla 156
Db 370 TTAATGGTATTAAGACAGATCTGATCATCA--ACACAGGTAAACAGAGAAAGAGA 426
Qy 157 GluLeuValIYsSerGlyAspGluAsnSerAlaIYrPheGluValSerAlaIYsAsn 176
Db 427 CAACAGTTA---GCACGCGACCTTAAGTTAACTATCATGAGGCGATCAGCAAGATTAGG 483
Qy 177 ThrAsnValAspGluMetPheTyrValLeuPheSerMetAlaIYsLeuProHisGluMet 196
Db 484 ATGAATGTAGATCAAGCTTTCATCAACTTGTCCGGTTATCAGAAATTTCAAGAGCAG 543
Qy 197 -----SerProAlaLeuHisArgIYs 203
Db 544 GAATGTCTCTCTTACCAAGAACCAACAGGAAA 576

```

## RESULT 9

```

PCT-US95-06420-5
Sequence 5, Application PC/TUS9506420
GENERAL INFORMATION:
APPLICANT: AARONSON, S.A.; CHAN, A.;
APPLICANT: MIKI, T.
TITLE OF INVENTION: NOVEL HUMAN RAS-RELATED
TITLE OF INVENTION: ONCOGENES UNMASKED BY EXPRESSION OF CDNA
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: MORGAN & PINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06420
FILING DATE: 24-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,946
FILING DATE: 24-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4150PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 615
TYPE: Nucleic acid

```

```

STRANDEDNESS: Double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: TC21 gene
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PCT-US95-06420-5
Alignment Scores:
Pred. No.: 1.36e-29 Length: 615
Score: 285.00 Matches: 74
Percent Similarity: 53.40% Conservative: 28
Best Local Similarity: 38.74% Mismatches: 75
Query Match: 20.70% Indels: 14
DB: Gaps: 4
US-09-778-963a-2 (1-266) x PCT-US95-06420-5 (1-615)

```

```

Qy 17 LysAsnSerYrArgMetValIleuGlyAlaSerArgValGlySerSerIleVal 36
Db 34 CAGAGAAATACCGCTGCTGTGTCGCGCGGCGCGCGTGGCAAGTCGCGCTACC 93
Qy 37 SerArgPheLeuAsnGlyArgPheGluAspGlnIYrThrProThrIleGluAspPheHis 56
Db 94 ATCCAGTTCATCCAGTCTCTTATTGTAACGATTATATGACACATTGAAGATTCTTAC 153
Qy 57 ArglyValIYrAenIleArgGlyAspMetYrGlnLeuAspIleLeuAspThrSerGly 76
Db 154 ACAAGAGCGTGTGATGATGACAGACAGACCCGGCTAGATATTGATGATACAGACGA 213
Qy 77 AsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeu 96
Db 214 NNNGAAGATTGGAGCATGAGAAACATATATAGAGCATGGCGAAGCTTCCTGTTG 273
Qy 97 ValPheSerLeuAspAsnArgGluSerPheAspGluValAsArgLeuGlnIle 116
Db 274 GTCTTTTCAGTACAGATGAGGCGAGTTTGAAGAACTTAACTTCAAGACAGATT 333
Qy 117 LeuGluValIYsSerCysLeuIYsAsnIYsThrIYsGluAlaIleGluLeuProMetVal 136
Db 334 CTCAGAGTA-----AAGATCGTGAAGATTTCCCAATGATT 369
Qy 137 IleCysGlyAsnIYsAsnAspHisGlyGluLeuCysArgIYsValProThrGluAla 156
Db 370 TTAATGGTATTAAGACAGATCTGATCATCA--ACACAGGTAAACAGAGAAAGAGA 426
Qy 157 GluLeuValIYsSerGlyAspGluAsnSerAlaIYrPheGluValSerAlaIYsAsn 176
Db 427 CAACAGTTA---GCACGCGACCTTAAGTTAACTATCATGAGGCGATCAGCAAGATTAGG 483
Qy 177 ThrAsnValAspGluMetPheTyrValLeuPheSerMetAlaIYsLeuProHisGluMet 196
Db 484 ATGAATGTAGATCAAGCTTTCATCAACTTGTCCGGTTATCAGAAATTTCAAGAGCAG 543
Qy 197 -----SerProAlaLeuHisArgIYs 203
Db 544 GAATGTCTCTCTTACCAAGAACCAACAGGAAA 576

```

## RESULT 10

```

US-09-778-963a-25

```

```

: Sequence 25 Application US/09765298A
: Patent No. 6582927
: GENERAL INFORMATION:
: APPLICANT: ARONHEIM, AMI
: TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A KIT THERE
: FILE REFERENCE: 108387.01
: CURRENT APPLICATION NUMBER: US/09/765, 298A
: CURRENT FILING DATE: 2001-01-22
: PRIOR APPLICATION NUMBER: IL 125456
: PRIOR FILING DATE: 1998-07-22
: PRIOR APPLICATION NUMBER: IL 128017
: PRIOR FILING DATE: 1999-01-12
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 25
: LENGTH: 551
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-765-298A-25

Alignment Scores:
Pred. No.: 4.73e-29 Length: 551
Score: 280.50 Matches: 67
Percent Similarity: 56.15% Conservative: 38
Best Local Similarity: 35.83% Mismatches: 63
Query Match: 20.37% Indels: 19
DB: 4 Gaps: 5

US-09-778-963A-2 (1-266) x US-09-765-298A-25 (1-551)
QY TyraAgmEtvAlValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPhe 39
D4 TATAAGCTGGTGGTGGTGGGCGCGCGCTGTGGCAAGTGGCGTGAACATCCAGCTG 63
QY LeuAengIyAArgPheGluAAspGlnTyrThrProThrIleGluAAspPheIleArgLysVal 59
D6 ATCCAGAACCACTTTGTGTGACGAAATACGACCCCACTATAGAGATTCCTACCGAAGAG 123
QY TyraenIlleArgIyAspMetTyrGlnLeuAAspIleLeuAAspThrSerGlyAsnHisPro 79
D124 GTGGCATGTGATGGGAGACGCTGCTGTGGACATCCTGATACCGCGCCGACGAGAG 183
QY PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 99
D184 TACAGCGGCATGCGGACCACTACATGCGACCGGAGGGCTTCGTGTGTGGCC 243
QY LeuAAspAAsnArgLysSerPheAAspGluValLysArgLeuGlnIleLeuGluVal 119
D244 ATCAACAACACCAAGTCTTTGAGAGACATCCACCACTACAGGACAGATC----- 294
QY LysSerCysLeuLysAsnLysThrLysGluAlaAlaGluLeuProMetValIleCysGly 139
D295 -----AAACGGGTGAAGGATCGGATGACGTGCTGATCGGCTGCTGCTGCTG 339
QY AsnLysAsnAAspHisGlyGluLeuCysArgGlnValProThrThrGluAlaGluLeu 159
D340 AACCAAGTGTGACTGGCT-----GCACGCACTGTGAATCTCGCAGAGCTCAGACCTC 393
QY ValSerGlyAspGluAAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnVal 179
D394 ---GCCCAAGCTACGCGCATCCCTCATCATCAAGACTCGGCCAAGACCCGCGAGAGT 450
QY AspGluMetPheTyrValLeu-----PheSerMetAlaLysLeu--- 192
D451 GAGATGCTTCTTACAGCTTGTTGCTGATGATCCGCGACGACAAAGCTCGGAACTGAAC 510
QY ---ProHisGluMetSerPro 198
D511 CCTCCTGATGAGAGTGGCCCC 531

RESULT 11
US-09-223-588-1
: Sequence 1, Application US/09223588

```

```

; Patent No. 6548261
; GENERAL INFORMATION:
; APPLICANT: Smith, Mark A.
; APPLICANT: McShea, Andrew
; TITLE OF INVENTION: Alzheimer Model for Drug Screening
; FILE REFERENCE: CASE-03597
; CURRENT APPLICATION NUMBER: US/09/223,588
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 1
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-223-588-1

Alignment Scores:
Pred.: 4.96e-29      Length: 567
Score: 280.50       Matches: 63
Percent Similarity: 58.68%    Conservative: 35
Best Local Similarity: 37.72%   Mismatches: 58
Query Match: 20.37%           Indels: 11
DB: 4                      Gaps: 3

US-09-778-963A-2 (1-266) x US-09-223-588-1 (1-567)

Oy      20 TyrArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPhe 39
Db      10 TATTAACCTGTGGTGGTAGTGAGACTGTGGCGGTGAGCAAGAGTCCTTGAAGATACAGCTA 69
Oy      40 LeuAsnGlyArgPheGluAspGlnTyrThrProThriIleGluAspPheHisArgLysVal 59
Db      70 ATTCAAGAATCATTTTGTGTGCAGCAATATGATCCAAACAATAGAGATTCTTACAGAACCAA 129
Oy      60 TyrAsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 79
Db      130 GTAGTAAATTGARGAGAAACCTGTCTTGGATATTCGCACACAGACAGGTCAAGAGAG 189
Oy      80 PheProAlaMetArgArgLysSerIleLeuThrGlyAspValPheIleLeuValPheSer 99
Db      190 TACAGGACATGAGGAGCACCAGTACATGACGAGCTGGGAGGAGCTTCTTGTGTAATTTGCC 249
Oy      100 LeuAspAsnArgGluSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluVal 119
Db      250 ATTAATAATTAATCTTAATCATTTGAAAGATATTCACATTATTAAGAAACAATTT----- 300
Oy      120 LysSerCysLeuLysAsnIleThrTyrGlnLalaIleGluLeuPrometValIleCysGly 139
Db      301 -----AAAAGACTTAAGACATCTGAAAGTACCTTAAGTGTCTTGAAGTACGA 345
Oy      140 AsnLysAsnAspHisGlyGlnLeuCysArgGlnValProThrThrGluAlaGluLeu 159
Db      346 AATAAATGATGATTTGGCTTCT-----AGAACAGTAGACACAAAAACAGGCTCAGACCTTA 399
Oy      160 ValSerGlyAspGluAsnSerIleTyrPheGluValSerIleAlaLysLysAsnThrAsnVal 179
Db      400 ---GCAGAAGATTAGCAATTCCTTTTATTGAACAATCAGCAAAAGACAGACAGGTTT 456
Oy      180 AspGluMetPheTyrValIleu 186
Db      457 GATGATGCTTCTTATACATTA 477

RESULT 12
US-08-884-866A-2
; Sequence 2, Application US/0884866A
; GENERAL INFORMATION:
; APPLICANT: Chien, Shu
; APPLICANT: Shyy, John Y-J
; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
; FILE REFERENCE: BYPASS
; FILE REFERENCE: UCSD1100-1
; CURRENT APPLICATION NUMBER: US/08/884,866A
; CURRENT FILING DATE: 1997-06-30

```

PRIOR APPLICATION NUMBER: 60/030,358  
PRIOR FILING DATE: 1996-11-08  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 570  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(570)  
US-08-884-866a-2

Alignment Scores:  
Pred. No.: 5e-29 Length: 570  
Score: 280.50 Matches: 67  
Percent Similarity: 56.68% Conservative: 39  
Best Local Similarity: 35.83% Mismatches: 62  
Query Match: 20.37% Indels: 19  
DB: Gaps: 5

US-09-778-963a-2 (1-266) x US-08-884-866a-2 (1-570)

Qy 20 TyrArgMetValIleuGlyAlaSerArgValGlySerSerIleValSerArgPhe 39  
10 TATAAGCTGGTGGTGGGCGCGCGCGTGGGCAAAAATGGCGTCAACATCAGCTG 69  
Qy 40 LeuAsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgVal 59  
70 ATCCAGAACCATTTTGTGGAGAAATACGACCCCACTAAGAGAGATTCTTCCGAGAGAG 129  
Qy 60 TyrAsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 79  
130 GTGGTCAATTATGGGAGAGAGCTGCTGTGGACATCTGTGATCCGCGCGCTGAGAGAG 189  
Qy 80 PheProIleMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 99  
190 TACAGCGCGCATGGGAGACAGTCAATGCGACCGGGAGGGCTTCTGTGTGTGGCC 249  
Qy 100 LeuAspAsnArgGluSerPheAspGluValIlyAsrGluGlnIlyGlnIleGluVal 119  
250 ATCAACAACCAAGTCTTTTGGAGACATCCACAGTACAGGAGAGCATC----- 300  
Qy 120 LysSerCysLeuLysAsnLysThrIleGluIleAlaGluLeuProMetValIleCysGly 139  
301 -----AAACGGGTAAAGACTCGAGTACGTCGCCATGCTGCTGTGGG 345  
Qy 140 AsnLysAsnAspHisGlyGluLeuGlySerGlnValProThrThrGluAlaGluLeu 159  
346 AACAAAGTGTACCTGGCT-----GCAAGCACTGTGAAATCTCGGACAGCTCAGGACTTC 399  
Qy 160 ValSerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIlyValYAsnThrAsnVal 179  
400 ---GCCCGAAGCTACGGCATCCCTACATCGAGACCTCGGCAAGACCGGACAGGAGTG 456  
Qy 180 AspGluMetPheTyrValIleu-----PheSerMetAlaIlyLeu--- 192  
457 GAGGATCCCTTACAGCTGTGTGCGTGAGATCCGGACAGCAAGCTCGGAACTGAAAC 516  
Qy 193 ---ProHisGluMetSerPro 198  
517 CCTCTGATGAGACTGGCCCC 537  
Db

RESULT 13  
US-09-167-322-12  
Sequence 12, Application US/09167322  
Patent No. 6365151  
GENERAL INFORMATION:  
APPLICANT: Allegheny University of the Health  
Sciences, Halpern, Michael S.  
TITLE OF INVENTION: CANCER VACCINE  
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P. C.  
STREET: Suite 1800, Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09167,322  
FILING DATE: 07-Oct-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/00582  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 7933-33 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4480 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-167-322-12  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Alignment Scores:  
Pred. No.: 1.46e-27 Length: 4480  
Score: 280.50 Matches: 68  
Percent Similarity: 54.92% Conservative: 38  
Best Local Similarity: 35.23% Mismatches: 68  
Query Match: 20.37% Indels: 19  
DB: Gaps: 5

US-09-778-963a-2 (1-266) x US-09-167-322-12 (1-4480)

Qy 14 ValProIleLysAsnSerTyrArgMetValIleuGlyAlaSerArgValGlyLysSer 33  
1248 GTAGAACGATGACAGAAATACCAAGCTTGTGTGGCGCTAGAGCGTGGAAAGGT 1307  
Qy 34 SerIleValSerArgPheLeuAsnGlyArgPheGluAspGlnTyrThrProThrIleGlu 53  
1308 GCCCTGACATCCAGCTGATCCAGAACCATTTGTGGACGATGTGATCCCATATGAG 1367  
Qy 54 AspPheHisArgLysValIlyAsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAsp 73  
1368 GACTCCACCGGAACAGATGATGATGAGGAGACGCTTACTGACATCTTAC 1427  
Qy 74 ThrSerGlyAsnHisProPheProIleMetArgArgLeuSerIleLeuThrGlyAspVal 93  
1428 ACAGCAGGTCAAGAAAGATATGATGCGGAGACAGTACATGCGCACAGGGAGGCGC 1487  
Qy 94 PheIleLeuValPheSerLeuAspAsnArgGluSerPheAspGluValIlyAsrGluGln 113  
1488 TTCTGTGTGATTTGCAATCAACACCAAGTCTTTTGAAGACATCCATCAGTACAG 1547  
Qy 114 LysGlnIleLeuGluValIlySerCysLeuLysAsnLysThrIlySerGluAlaIleGlu 133  
1548 GAGCAGATC-----AAGCGGGAAGAAATTCAGATGAGTGTG 1583  
Qy 134 ProMetValIleCysGlyAsnLysAsnAspHisGlyGluLeuGlyAsrGlnValProThr 153  
1584 CCAATGTGTGTGGTGGGCAACAGTGTGACCTGGCC-----GCTCACACTGTGAGTCT 1637  
Db



```
Db      202 TATAAACTTGGTGTGAGCTGTGCGGTAGGCAAGAGCTTACGATACAGCTA 261
Qy      40 LeuAenGIyArGPheGIuAspGIInTyThrProThrIIeGIuAspPheHISArGIyVal 59
Db      262 ATTCAAGATCATTTGTGTGAGCAATATGATCCACATATAGAGATTCTTACAGAAAGCA 321
Qy      60 TyrAsnIIeArGIyAspMetTyrGIInLeuAspIIeLeuAspThrSerGIyAsnHISPro 79
Db      322 GTAGTAATTGATGGAGAAACCTGCTCTTGGATATTCTGACACAGCAGGTCAAGAGAG 381
Qy      80 PheProAlaMetArGIyArGIleuSerIIeLeuThrGIyAspValPheIIeLeuValPheSer 99
Db      382 TACAGTCGAATGAGGAGCACGATACATGAGACTGGGAGGGCTTCTTGTGTATTGCGC 441
Qy      100 LeuAepAenArGIuSerPheAspGIuValIySArGIleuGIInIIeLeuGIuVal 119
Db      442 ATAAATATACTTAATCATTTGAAAGATATTCACCATATAGAGAACAAATT----- 492
Qy      120 LysSerCysLeuIyAsnIySArGIyGIuAlaIaGIuLeuProMetValIIeCysGIy 139
Db      493 -----AAAGAGTTAAGACTCTGAAAGATGTAAGTACTATAGTCTTACTAGGA 537
Qy      140 AsnIySAenAspHISGIyGIuLeuCysArGIInValProThrThrGIuAlaGIuLeu 159
Db      538 AATAATGTGATTTGCCCTCT-----AGAACAGTAGACACAAACAGGCTCAGACTTA 591
Qy      160 ValSerGIyAspGIuAsnSerAlaTyrPheGIuValSerAlaIyGIyAsnThrAsnVal 179
Db      592 ---GCAAGAACTTATGCAATTCCTTTATGTGAACATCAGCAAGAACAAAGACAGGCTGT 648
Qy      180 AspGIuMetPheTyrValIeu 186
Db      649 GATGATGCTTCTATACATTA 669
```

Search completed: June 20, 2004, 22:59:30  
Job time : 94 secs



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: June 20, 2004, 20:24:43 ; Search time 2770 Seconds

(without alignments)  
2867.632 Million cell updates/sec

Title: US-09-778-963A-2  
Perfect score: 1377  
Sequence: 1 MMKTLSSGNCITLSPVANKSY.....IKAKVLRGQARERDKCTIQ 266

Scoring table: BIOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Rgapop 6.0 , Rgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xjh  
-Q=/cgm2\_1/USPRO.spool/US09778963/runat\_18062004\_123500\_15907/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biom62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09778963 @CGN\_1\_1\_3437 @runat\_18062004\_123500\_15907 -NCPU=6 -ICPU=3  
-NO MAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1347	97.8	1006	13 BQ073742	BQ073742 AGENCOURT
2	1320	95.9	947	13 BQ954076	BQ954076 AGENCOURT
3	1304	94.7	1300	11 AK015898	AK015898 Mus muscu
4	1299	94.3	904	13 BQ947936	BQ947936 AGENCOURT
5	1206.5	87.6	888	12 B1754083	B1754083 603027639
6	1184	86.0	879	13 BX704351	BX704351 BX704351
7	935	67.9	808	13 BU208822	BU208822 603949137
8	910	66.1	886	13 BX704490	BX704490 BX704490
9	879	63.8	1780	11 BC042688	BC042688 Homo sapi
10	874.5	63.5	1474	11 AK038932	AK038932 Mus muscu
11	869.5	63.1	761	14 CF999619	CF999619 AGENCOURT
12	859	62.4	891	13 BX327655	BX327655 BX327655
13	853.5	62.0	726	14 CF997752	CF997752 AGENCOURT
14	829	60.2	834	14 CD513046	CD513046 AGENCOURT
15	807	58.6	749	14 CK029404	CK029404 AGENCOURT
16	801.5	58.2	809	13 BX880689	BX880689 BX880689
17	797.5	57.9	906	13 BU909336	BU909336 AGENCOURT
18	797.5	57.9	991	14 CB205954	CB205954 AGENCOURT
19	796	57.8	815	12 BG966663	BG966663 602834520
20	789	57.3	848	14 CF375942	CF375942 AGENCOURT
21	785	57.0	751	14 CF522567	CF522567 AGENCOURT
22	782	56.8	822	14 CD251380	CD251380 AGENCOURT
23	781.5	56.8	985	14 CB205910	CB205910 AGENCOURT
24	781	56.7	1035	12 BM919341	BM919341 AGENCOURT
25	776	56.4	628	12 BG713443	BG713443 pg11n.pko
26	774.5	56.2	762	14 CA367297	CA367297 pg11n.pko
27	772.5	56.1	707	14 CA355110	CA355110 626946 NC
28	771.5	56.0	726	12 B1622707	B1622707 B1622707
29	771.5	56.0	809	10 BE408863	BE408863 601303729
30	768	55.8	962	12 BM543472	BM543472 AGENCOURT
31	760	55.2	914	14 CD513372	CD513372 AGENCOURT
32	757.5	55.0	740	14 CF547394	CF547394 AGENCOURT
33	754.5	54.8	726	10 BF613135	BF613135 de30e03.Y
34	753	54.7	742	9 AM028127	AM028127 wv2c07.X
35	752.5	54.6	697	14 CA355132	CA355132 626970 NC
36	748	54.3	618	12 BJ069624	BJ069624 BJ069624
37	745.5	54.1	710	12 BJ526038	BJ526038 BJ526038
38	745	54.1	855	14 CD105429	CD105429 AGENCOURT
39	742	53.9	958	13 BQ719566	BQ719566 AGENCOURT
40	730	53.0	627	10 BE390509	BE390509 601284127
41	725.5	52.7	640	12 BJ490945	BJ490945 BJ490945
42	723.5	52.5	1103	12 BM920514	BM920514 AGENCOURT
43	723	52.5	655	12 BG711792	BG711792 pg11n.pko
44	716.5	52.0	1053	12 BM921737	BM921737 AGENCOURT
45	713	51.8	626	10 BE729219	BE729219 601560842

#### ALIGNMENTS

RESULT 1  
LOCUS BQ073742 1006 bp mRNA linear EST 02-APR-2002  
DEFINITION AGENCOURT 7046611 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:5806628  
5', mRNA sequence.  
ACCESSION BQ073742  
VERSION BQ073742  
KEYWORDS EST, sapiens (human)  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1006)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
 Plate: LNCM2047 row: 1 column: 21  
 High quality sequence stop: 671.

## FEATURES

source  
 1..1006  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:580628"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_101"  
 /note="Organ: lung; Vector: pOT87; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,62e-163 Length: 1006  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 98.87% Conservative: 2  
 Best Local Similarity: 98.11% Mismatches: 3  
 Query Match: 97.82% Indels: 0  
 DB: 13 Gaps: 0

US-09-778-963a-2 (1-266) x B0073742 (1-1006)

QY 1 MetcetyThrLeuSerSerGlyAsnGlyThrLeuSerValProAlaIysAsnSerTyr 20  
 DB 29 ATGATGAAGACCTTGTCCAGCGGAACTCCAGCTCAGTGTCCCGCAAAACTCATTAC 88  
 QY 21 ArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40  
 DB 89 CCGCATGTGTCTCTGGTCTCTCTGCGGGTGGCAAGAGCTCATGTCTGCTTCTCTC 148  
 QY 41 AangIyArgPheGlyAspGlnTyrThrProThrIleGlyAspPheHisArgIyValTyr 60  
 DB 149 AATGCGCGCTTGAAGACCACTACACACCACCATGAGACTTCCACCGTAAGGTATAC 208  
 QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 DB 209 AACATCCGGCGGACATGTACAGCTCCAGATCTCGATCTCGGCAACACCCCTTC 268  
 QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 DB 269 CCGGCATGCGAGGCTGTCTCTACAGGGAGTGTCTTCATCTGTGTTTCAGCTCG 328  
 QY 101 AspAsnArgGlyLeuSerPheAspGlnValLysArgLeuGlnLysGlnIleLeuGlnValLys 120  
 DB 329 GATTAACCGGAGCTCTTGCATAGAGTCAAGCCCTTCAAGACACATCTCGAGGTCAAG 388  
 QY 121 SerCysLeuLysAsnLysThrIyGlnAlaIaIaGlnLeuProMetValIleCysGlyAsn 140  
 DB 389 TCCTGCTGTAAGAACCAAGACCAAGAGCGGCGAGCTGCCCATGTCTATCTGTGGCAC 448  
 QY 141 LysAsnAspHisGlyGlnLeuCyArgGlnValProThrThrGlnAlaGlnLeuLeuVal 160

DB 449 AAGAACGACCAAGCGGAGCTGTGCGCCGACGTTGCCACCAACCGAGCCGAGCTGTGTC 508  
 QY 161 SerGlyAspGluAsnSerAlaTyrPheGlnValSerAlaLysAsnThrAsnValAsp 180  
 DB 509 TCGGCGCAGACGAACTGCGGCTTACTTCCAGAGGTGTGGCCAAAGAACACCAACGTTGAC 568  
 QY 181 GJMetPheTyrValLeuPheSerMetAlaLysLeuProHisGlnMetSerProAlaLeu 200  
 DB 569 GAGATGTTCTAGTCTCTTCAAGCTGGCCAAAGCTGCCACAGAGATAGGCCCGCTTG 628  
 QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
 DB 629 CATGCCAAGATCTCGTCCAGTACAGTACGCCCTTCCACCCAGGCCCTTCTGCATGGCC 688  
 QY 221 ArgValLysGlnMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240  
 DB 689 CCGCTCAAGAGATGACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 748  
 QY 241 AsnSerAspLeuLysTyrIleValAlaLysValLeuArgGlnGlnAlaArgGlnArg 260  
 DB 749 AACAGTGAACCTTCAATCATCAAGGCAAGGATCTTGGGAAGGCCAGCCGTGAGAGAG 808  
 QY 261 AspLysCysThrIle 265  
 DB 809 GACAAATGCACCATC 823

RESULT 2  
 B0954076 947 bp mRNA linear EST 21-AUG-2002  
 LOCUS AGNCOURT 8866147 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6464135  
 DEFINITION 5', mRNA sequence.

ACCESSION B0954076 GI:22369554  
 VERSION B0954076  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 947)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
 Plate: LNCM2656 row: 1 column: 24  
 High quality sequence stop: 668.

## FEATURES

## source

1..947  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6464135"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_101"  
 /note="Organ: lung; Vector: pOT87; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

## ORIGIN

## Alignment Scores:





## Alignment Scores:

Pred. No.: 2,34e-157 Length: 904  
 Score: 1299.00 Matches: 260  
 Percent Similarity: 98.13% Conservative: 3  
 Best Local Similarity: 97.01% Mismatches: 3  
 Query Match: 94.34% Indels: 3  
 Gaps: 0

US-09-778-963a-2 (1-266) x B0947936 (1-904)

```

QY 1 MetMeclyThrLeuSerSerGlyYasnCyThrLeuSerValProAlLysAsnSerTyr 20
DB 29 ATGATGAAGAACTTTCTCCAGCGGAACTGCACAGCTCAGATGTCGCCGCAAAATCATATC 88
QY 21 ArgMetValLleuGlyAlaSerArgValGlyYsserSerLeuValSerArgPheLeu 40
DB 89 CGCATGTGTGTCTCGGTGCTCTCGGTGGGCAAGAGCTCATGCTGTCTGCTTCTC 148
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgIleValTyr 60
DB 149 AATGSCCGCTTTGAGACCAAGTACACACCCACATCGAGACTTCACCGTAAAGGTATAC 208
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyYasnHisProPhe 80
DB 209 AACATCCCGCGGACATGTATACAGCTCCAGATCCTCGATACCTCTGGCAACACCCCTTC 268
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
DB 269 CCCGCGATGCGGACGCTCTCCATCTCTCACAGGGAGTCTTCATCTGCTGTTCAGCTCG 328
QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnIleLeuGluValLys 120
DB 329 GATTAACCGGAGTCTTGTGATGAGGTCAAGCCCTTCAGAACCAATCTCGAGGTCAAG 388
QY 121 SerCysLeuYsnAsnIleThrLysGluAlaIleGluLeuProMetValIleCysGlyAsn 140
DB 389 TCTGCTCCGAGAAACAAGACCAAGAGGCGGCGAGCTGCCATGTATCTGTGGCAAC 448
QY 141 LysAsnAspPheIleGlyLeuLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160
DB 449 AAGAACGACCAACGCGAGCTGTGCGGCAAGGTGCCACACCAAGGCGGAGCTGTGAGT 508
QY 161 SerGlyAspGluAsnSerValTyrPheGluValSerAlaValLysAsnThrAsnValSer 180
DB 509 TCGGCGACGAGAACTGCGCTTACTTCAGAGGTGTGCGCAAGAAACCAACGCTGAC 568
QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
DB 569 GAGATGTTCTACGTCTCTTCAGCATGCGCAAGCTGCCACAGATGAGAGCCCGCTCG 628
QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220
DB 629 CATGCAAGATCTCGGTGAGTACGAGTACGCGCTTCACACCGGCGCTTGTGATGAGC 688
QY 221 Arg-ValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240
DB 689 CGCGGTAAAGGAGATGAGCGCTATGCAATGCTGCGCTTGCAGCGCGCGCGAGGCT 748
QY 240 LysSerAspLeuLysTyrIleLysAlaLysValLeu-ArgGluGlnIleAlaTyrGlu 260
DB 749 CAA-AGTAACTCAAGTACATCAAGCCAAAGTCTTCGCGGAAAGCCAGGCGGTGAGCA 807
QY 260 rgaApLysCysThrIleGln 266
DB 808 AGGACAGTGCACCATCAA 827

```

RESULT 5  
 LOCUS B1754083 888 bp mRNA linear EST 25-SEP-2001  
 DEFINITION 603027639p1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5198024 5',  
 mRNA sequence.  
 ACCESSION B1754083

## VERSION B1754083.1 GI:15745661

## KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

1 (bases 1 to 888)

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Contact: Robert Strassberg, Ph.D.  
 Email: cgsapb@remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LIML at:  
<http://image.liml.gov>  
 Plate: LIML1495 row: n column: 09  
 High quality sequence stop: 796.  
 Location/Qualifiers

## FEATURES

## source

1..888  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5198024"  
 /lab\_host="DH10B"  
 /clone\_id="NIH\_MGC\_114"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 019. Note:  
 this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,33e-145 Length: 888  
 Score: 1206.50 Matches: 245  
 Percent Similarity: 93.63% Conservative: 5  
 Best Local Similarity: 91.76% Mismatches: 9  
 Query Match: 87.62% Indels: 8  
 Gaps: 2

US-09-778-963a-2 (1-266) x B1754083 (1-888)

```

QY 1 MetMeclyThrLeuSerSerGlyYasnCyThrLeuSerValProAlLysAsnSerTyr 20
DB 90 ATGATGAAGAACTTTCTCCAGCGGAACTGCACAGCTCAGATGTCGCCGCAAAATCATATC 149
QY 21 ArgMetValLleuGlyAlaSerArgValGlyYsserSerLeuValSerArgPheLeu 40
DB 150 CGCATGTGTGTCTCGGTGCTCTCGGTGGGCAAGAGCTCATGCTGTCTGCTTCTC 209
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgIleValTyr 60
DB 210 AATGSCCGCTTTGAGACCAAGTACACACCCACATCGAGACTTCACCGTAAAGGTATAC 269
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyYasnHisProPhe 80
DB 270 AACATCCCGCGGACATGTATACAGCTCCAGATCCTCGATACCTCTGGCAACACCCCTTC 329
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
DB 330 CCCGCGATGCGGACGCTCTCCATCTCTCACAGGGAGTCTTCATCTGCTGTTCAGCTCG 389
QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnIleLeuGluValLys 120
DB 390 GATTAACCGGAGTCTTGTGATGAGGTCAAGCCCTTCAGAACCAATCTCGAGGTCAAG 449

```

QY 121 SerCysLeuIysAsnIysThrIysGluAlaIaGluLeuProMetValIleCysGlyAsn 140  
 Db 450 TCTGCTGTAAGAAACAGACAGAGAGCGCGGAGCTGCTGCTATCTGTGGCAAC 509  
 QY 141 LysAsnAspHisGlyGluLeuCysArgIleValProThrThrGluIaGluLeuVal 160  
 Db 510 AAGAAACGACACCGGAGCTGTGCTCCGAGGTCCACACAGAGCCGAGCTGTGG 569  
 QY 161 SerGlyAspGluAsnSerIaIysPheGluValSerAlaIysAsnThrAsnValAsp 180  
 Db 570 TCGGGGACAGAGACGCGCTTACCTTGAGGTGCGCCAGAAAGAACCAACGCGGAC 629  
 QY 181 GluMetPheTyrValLeuPheSerMetAlaIysLeuProHisGluMetSerProAlaLeu 200  
 Db 630 GAGATGTTTACGCTGCTTTCAGCATGCGCAAGCTCCACACAGATGAGCCCCGCTG 689  
 QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPhe-CysMetAr 220  
 Db 690 CATCGCAAGATCTCCGTGACGTACGCTTCCACCCGAGCCCTTCTTGATGCG 749  
 QY 220 G-ArgValIysGluMetAsp-AlaTyrGlyMetValSerProPheAlaArgArProSer 239  
 Db 750 CCGGCTCAAGAGATGAGACCGCTATGCGATGCTCGCCCTT-----CGCCGCGG 803  
 QY 240 ValAsnSerAsp-----LeuIysTyrIleIysAlaIysValLeuArgGluGlyGln 256  
 Db 804 GCCAGCGCTCAACAGAGTACCTCAGGTCCATTCAAGCCCAAGGTCTTGGGAAAGCGG 863  
 QY 257 AlaArgGluArgAsp 261  
 Db 864 GGCCGTGAGAGGAGC 878

RESULT 6  
 LOCUS BX704351 879 bp mRNA linear EST 17-NOV-2003  
 DEFINITION BX704351 XGC-tadpole silurana tropicalis cdna clone TTPA010e21 5',  
 mRNA sequence.  
 ACCESSION BX704351  
 VERSION BX704351.1 GI:38366558  
 KEYWORDS EST.  
 SOURCE Silurana tropicalis (western clawed frog)  
 ORGANISM Silurana tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Silurana.  
 REFERENCE Croining,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
 AUTHORS Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
 TITLE Unpublished (2003)  
 JOURNAL Contact: Croining MDR  
 COMMENT Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE\_ID: TTPA010e21.plkSP6  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Nigel Garrett.  
 cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with  
 Vector: pCS107; Site\_1: EcoRI; Site\_2: NotI  
 Host: Escherichia coli DH10B.  
 Location/Qualifiers  
 1..879  
 /organism="Silurana tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="TTPA010e21"  
 /dev\_stage="tadpole (stage 35-40)"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="XGC-tadpole"

## FEATURES

source  
 1..879  
 /organism="Silurana tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="TTPA010e21"  
 /dev\_stage="tadpole (stage 35-40)"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="XGC-tadpole"

/note="Vector: pCS107; Site\_1: EcoRI; Site\_2: NotI; cDNA  
 was oligo dt primed from 5ug of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pCS107  
 with EcoRI at the 5' end and NotI at the 3' end"

## ORIGIN

## Alignment Scores:

Pred. No.:	1..92e-142	Length:	879
Score:	1184.00	Matches:	228
Percent Similarity:	92.86%	Conservative:	19
Best Local Similarity:	85.71%	Mismatches:	1
Query Match:	85.98%	Indels:	0
DB:	13	Gaps:	0

US-09-778-963a-2 (1-266) x BX704351 (1-879)

QY 1 MetCysIysThrLeuSerSerGlyAsnCysThrIleuSerValProAlaIysAsnSerTyr 20  
 Db 57 ATGATGAAGACTCTCTCCAGTGGGAACTGCACACTCAGTGTCCAGCTAAACCTTAT 116  
 QY 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerArgPheLeu 40  
 Db 117 CGGATGTTGCTTCCTGGGGCATCCAGATTGGCAAGAGTGCATCGTTCTCGTTCTTG 176  
 QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60  
 Db 177 AACGAGCATTTGAAAGACCAATACACACCACTTGAAGATTTCATCGCAAGCTGAC 236  
 QY 61 AsnIleArgGlyAspMetTyrGlnIleuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 Db 237 AACATCAAGAGAGACATGTACAGCTGACATTCATGATACATCTGGAACCATTCATT 296  
 QY 81 ProAlaMetArgArgLeuSerIleuThrGlyAspValPheIleLeuValPheSerIleu 100  
 Db 237 CCAGCATGAGAGAGGCTATTCATTTCTACAGATGATGTTTCATCTTGTTCAGCTG 356  
 QY 101 AspAsnArgGluSerPheAspGluValIysArgLeuGlnIleuGluValIys 120  
 Db 357 GACAAACGGGACTCCTTGATGAAAGTAAAGCGCTTGAAGCAAAATTTGAAAGTCAA 416  
 QY 121 SerCysLeuIysAsnIysThrIysGluAlaIaGluLeuProMetValIleCysGlyAsn 140  
 Db 417 TCTGTGTCAAAAACAAAACTAAAGACACAGTGAATTCCTATGATGATTTGTGGCAAC 476  
 QY 141 LysAsnAspHisGlyGluLeuCysArgGluValProThrThrGluIaGluLeuVal 160  
 Db 477 AAAGTGTATCTAGGGAGACATCACCGCAGGTGAGAGCTGAAGAGAGCGCTGGTG 536  
 QY 161 SerGlyAspGluAsnSerIaIysPheGluValSerAlaIysIysAsnThrAsnValAsp 180  
 Db 537 TCTGGAGATGAACCTGTGCTTACCTTGAGATTTCTGCCAAAGAAATATTAAGTAC 596  
 QY 181 GluMetPheTyrValLeuPheSerMetAlaIysLeuProHisGluMetSerProAlaLeu 200  
 Db 597 AAAATGTTCCAAAGCTCAATTAGCATGCAAAAGCTGCCAATGAAGAGCCGACCTC 656  
 QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
 Db 657 CATGGAAGATCTCCATGACGATATGGGATACCTTTACACGAATCTTTCAGGCTAGA 716  
 QY 221 ArgValIysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArProSerVal 240  
 Db 717 AGAATGAAGAGAGAGATGATGCTTATGAGATGTCTCACTTCAGTAAACGTCG 776  
 QY 241 AsnSerAspLeuIysTyrIleIysAlaIysValLeuArgGluGlyGlnAlaArgGluArg 260  
 Db 777 AACAGTACCTTAAGTACATCAAGTCAAAAGTACTGGGGAAGAGACAGGACGAG 836  
 QY 261 AspIysCysThrIleGln 266  
 Db 837 GAGAACTGCACTATTCAAG 854

RESULT 7



BU208822 808 bp mRNA linear EST 25-NOV-2002  
 LOCUS 603949137F1 CSECHN03 Gallus gallus cDNA clone ChEST906j10 5', mRNA  
 DEFINITION  
 ACCESSION BU208822  
 VERSION BU208822.1 GI:25380427  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 808)  
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 TITLE A Comprehensive Collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 2235534  
 PUBMED 12445392  
 COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QP, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1..808  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hisex"  
 /db\_xref="taxon:9031"  
 /clone="CHEST906j10"  
 /tissue\_type="whole embryo"  
 /dev\_stage="20-21"  
 /lab\_host="DH10B"  
 /clone\_lib="CSECHN03"  
 /note="Organ: whole embryo; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones; cDNA synthesis was initiated using an oligo(dt) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996) 791, except that a significantly longer reannealing hybridization was used."

ORIGIN  
 Alignment Scores: 3,61e-110 Length: 808  
 Pred. No.: 935.00 Matches: 197  
 Score: 92.44% Conservative: 11  
 Best Local Similarity: 87.56% Mismatches: 16  
 Query Match: 67.90% Indels: 4  
 Gaps: 0  
 DB: 13  
 US-09-778-963a-2 (1-266) x BU208822 (1-808)

QY 1 MetMetlySThrLeuSerSerGlyAsnCySThrLeuSerValProAlaIlyAsnSerTyr 20  
 Db 133 ATGATGAGAACCACTGCTAGTGAAGAACTGCACCTGAGACGTCGACGACAGACTCGTAC 192  
 QY 21 ArgMetValIleLeuGlyAlaSerArgValGlyIlySerSerIleValSerArgPheLeu 40  
 Db 193 CGCATGGGTGGTGGAGCTCCAGAGGTGGGAAATTCCTCATAGTCTCAGCGCTTTC 252  
 QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheNleArgIlyValTyr 60

Db 253 AATGGCCGGTTTGAGAGCACTACACCTCCACCAATTGAGATTTTCAACCGCAAGCTTAC 312  
 QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 Db 313 AACATCCGGGAGATATGTATACCTGACCTGGACATCTCCGACACCTCAGGGAATCACCTTTC 372  
 QY 81 ProAlaMetArgAlyLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 Db 373 CTGCTATGAGAGAGCTTTTCCATCTTCAGAGCCAGATGTTTCACTCTGGATTTCAGCTTG 432  
 QY 101 AspAsnArgGluSerPheAspGluValIlyAspArgLeuGlnIleLeuGluValIly 120  
 Db 433 GATAACAGAGATATCTTTGATGAGAGTCAAGAGACTCCAGAGACGATCTCGAGGTCAA 492  
 QY 121 SerCysLeuIlyAsnIlySThrIlyGluAlaIleGluLeuProMetValIleCysGlyAsn 140  
 Db 493 TCTGCTCGAAGAACAGACCAAGAAATCACTCACTCCCATGCTGATCTGTGGCAAC 552  
 QY 141 IlyAsnAspHisGlyGluLeuCyArgGlnValProThrThrGluAlaGluLeuVal 160  
 Db 553 AAAATGACCAAGTGAATATACCGCAAGTGGCTCAGATGAAGTGAGAACTTGTG 612  
 QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIlyAsnIlyAsnValAsp 180  
 Db 613 TCCAGTATGAAACTGCGCTTACTTCAAGATTTCAGTAAGAAACACCACTGAT 672  
 QY 181 GluMetPheTyrValLeuPheSerMetValIlyLeuProHisGluMetSerProAlaLeu 200  
 Db 673 GAGATGTC-TATGTCTCTTCAGATGCGCAAGCTACT-CATGAGATGAGGCCCTGCCCCCTC 730  
 QY 201 HisArgIlyIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCyMetAr 220  
 Db 731 CAACAGGAAGACTTCATCTCAGTACGAGGAGACACCTTCCACAGAGAATC-TTCCGATGCG 789  
 QY 220 GARGValIlyGlu 224  
 Db 790 ACGAGTGAAGAC 802

RESULT 8  
 EX704490  
 LOCUS  
 DEFINITION BX704490 XGC-tadpole Silurana tropicalis cDNA clone TTPA008h19 5',  
 mRNA sequence.  
 ACCESSION BX704490  
 VERSION BX704490.1 GI:38366697  
 KEYWORDS EST.  
 SOURCE Silurana tropicalis (western clawed frog)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 Xenopodinae; Silurana.  
 REFERENCE 1 (bases 1 to 886)  
 AUTHORS Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M., and Rogers, J.  
 TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Croning MDR  
 Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE\_ID: TTPA008h19.plk56  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Nigel Garrett.  
 cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole  
 embryos. EcoRI-NotI cut cDNA was then ligated into pBS107 with  
 EcoRI at the 5' end and NotI at the 3' end.  
 Vector: pBS107; Site\_1: EcoRI; Site\_2: NotI  
 Host: Escherichia coli DH10B.  
 Location/Qualifiers  
 1..886  
 /organism="Silurana tropicalis"

FEATURES  
 source



/clone.lib="NH MGC\_114"  
/lab host="MDH10B"  
/note="Vector: pCMV-SPORT6"

## ORIGIN

## Alignment Scores:

Score: 2.28e-102 Length: 1780  
Percent Similarity: 879.00 Matches: 172  
Best Local Similarity: 77.70% Conservative: 44  
Query Match: 61.87% Mismatches: 46  
Indels: 16  
Gaps: 4

US-09-778-963a-2 (1-266) x BC042688 (1-1780)

```

Qy 1 MetcetythrleuserserglyasnCythrleuserValProAlaYasnSerTyr 20
Db 196 ATGATCAAGAGATGTGCGCCGAGGAGACTCGAGCTGATATCCGCGCAAGACTGCTAT 255
Qy 21 ArgmetValValleuGlyAlaSerArgValGlyYsserSerIleValSerArgPheleu 40
Db 256 CGCATGTGATCTCTGCTGCTCAAGGTGGCAAGAGGCGCATCTGTCTCGCTCTCTC 315
Qy 41 AsnGlyArgPheGlyAspGlnTyrThrProThrIleGlyAspPheHisArgLyValTyr 60
Db 316 ACCGCGCGCTTCGAGAGAGCTTACACGCTTACATCCAGACACTTCCACCGCAAGTCTAC 375
Qy 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleleuAspThrSerGlyAsnHisProPhe 80
Db 376 TCATTCGCGCGGCGAGGCTTACCATCGATCTCCGACATCTCCGCGGACACACCCGCTTC 435
Qy 81 ProAlaMetArgArgLeuSerIleleuThrGlyAspValPheIleleuValPheSerIleu 100
Db 436 CCCGCCATGCGCGCTCTCCATCTCAAGAGAGCGTTTCATCTCGTGTTCAGTCTG 495
Qy 101 AspaAsnArgGlySerPheAspGlyValIleArgLeuGlnIleLeuGlyValIle 120
Db 496 GACACCGCGCATCTCTTCGAGAGAGTGCAGCGCTCAGGACAGCATCTCGACACCAAG 555
Qy 121 SerCysLeuYsaenLyThrIleGlyAlaIleGlyLeuProMetValIleCysGlyAsn 140
Db 556 TCTTGCTCAAGAACAAACCAAGAGACGTGAGCGGCGCTGATCTGCGGCAAC 615
Qy 141 LysAsnAspHisGlyGlnLeuCyArgGlnValProThrThrGlnAlaGlnLeuVal 160
Db 616 AAGGATGACCGC--GACTTCTACCGGAGGTGACGACGCGAGATCGACACTGTGTG 672
Qy 161 SerGlyAsp--GluAsnSerAlaTyrPheGlyValSerAlaIleYsaenThrAsnVal 179
Db 673 GCGCAGACACCCCGCGCTGCGCTTACTTCAGATCTGCGCCAGAGACAGACAGCCTG 732
Qy 180 AspGlyMetPheTyrValLeuPheSerMetAlaIleYsaenProHisGlyMetSerProAla 199
Db 733 GACCAAGATGTTCCGCGCGCTCTTCGCGCATGCGCAAGCTGCGCAGAGATGAGCCCAAC 792
Qy 200 LeuHisArgLyIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMet 219
Db 793 CTGACCGCAGAGGCTCGTGTGACAGTGCACAGCTGTGCACAGAAAG-----GCGCTG 846
Qy 220 ArgArgValIleYsaenGluMet-----AspAla 227
Db 847 CGGAACAAGAGCTGCTGCGCGCGGCGGACGCGCGCGCGCGGACCGCGGAGAGCGCC 906
Qy 228 TyrGlyMetValSerProPheAlaArgArgProSerValIaenSerAspLeuLyThrIle 247
Db 907 TTGGCATCTGCGGACCTTCGCGCGCGCGCGCGGCGGACGAGACAGGACCTCATGACATC 966
Qy 248 LysAlaIleValIleuArgGlnGlyGlnAlaIleArgGlnArgAspLyGlyThrIle 265
Db 967 CGCAGAGAGGCGCAGCGCGGCGGACGAGCGCAAGAGCAAGAGGCGCTGCTGATTC 1020

```

## LOCUS

AK038932 1474 bp mRNA linear HTC 19-SEP-2003  
Mus musculus adult male hypothalamus cDNA, RIKEN full-length  
enriched library, clone:A230076M10 product:ras-related protein  
(DEXRAS1) mRNA, full insert sequence.

## ACCESSION

AK038932  
AK038932.1 GI:26332910

## VERSION

HTC; CAP trapper.

## KEYWORDS

Mus musculus (house mouse)

## SOURCE

Mus musculus

## ORGANISM

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

## TITLE

99279253

## JOURNAL

10349636

## MEDLINE

2

## PUBMED

3

## AUTHORS

Shibata,K., Itoh,M., Aizawa,K., Nagaoke,S., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,  
Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,K., Ohara,E., Matshiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

## TITLE

20530913

## JOURNAL

11076861

## MEDLINE

4

## PUBMED

5

## AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

## TITLE

6

## JOURNAL

6 (bases 1 to 1474)

## MEDLINE

Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

## PUBMED

6 (bases 1 to 1474)

## AUTHORS

Adachi,U., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
Hori,F., Imocant,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numasaki,K., Ohno,M., Ohnoto,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shitaki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.

## TITLE

Direct Submission

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
URL:htp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>.

## FEATURES

## SOURCE

Location/Qualifiers

1. 1474  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM DB:A230076M10"  
/db\_xref="MGI:2403379"  
/db\_xref="taxon:10090"  
/clone="A230076M10"  
/sex="male"  
/tissue\_type="hypothalamus"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
141..983  
/note="unnamed protein product; putative ras-related protein (DEKAS1) mRNA (GS|Af009246, evidence: BLASTN, 99%, match=1464)"  
/codon\_start=1  
/protein\_id="BAC30173.1"  
/db\_xref="GI:26332911"  
/translation="MKLAAMTKKCPSPSELSIPAKNCRVWILGSSVCKTATVSR LTRPREDAYPTIEDFHRKFSIRGEVYQILIDVDSNPPPMKRLSLTGDVFLV FSLDRDSEFEEVORLKOQILDTKSLCNKTKEMDVPLVIGNKNDPRFVREVEI EQLVDDPQRCAYFEISAKNSLSDMPRALFAMKLPSEMSPDILHRVVOYCVLH KVALRNKTLNRAGSGGGDHGDAFGILAPFARFVSHDLNVIKRETSVGSQAKDER CVIS"

## CDS

## ORIGIN

## Alignment Scores:

Pred. No.: 6,54e-102 Length: 1474  
Score: 874.50 Matches: 171  
Percent Similarity: 77.98% Conservative: 45  
Best Local Similarity: 61.73% Mismatches: 46  
Query Match: 63.51% Indels: 15  
Gaps: 4

US-09-778-963A-2 (1-266) x AK038932 (1-1474)

QY 1 MetMetLysThrLeuSerSerLysCysThrLeuSerValProAlaLysAsnSerTyr 20  
Db 156 ATGATCAAGAGATGTGCCCAAGCACTGTGAATGATCCGCGCAAGACTGTCTAC 215  
QY 21 ArgMetValValLeuGlyValAsnArgValGlyLysSerSerLysValSerArgPheLeu 40  
Db 216 AGGATGCTCATCTCGGCTCATCCCAAGTGGCCAGACGCCCATGTGTGCGGCTTCTC 275  
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrLysGluAspPheHisArgLysValTyr 60  
Db 276 ACGGCGCGTTTGAAGATGCTTACACCCCTCATCAGACTCCACCGAAGATTTC 335  
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
Db 336 TCGATCCGCGGGAAGTTCACAGTTGACATACATCCGACATCCGCGCATATCCGTTT 395  
QY 81 ProAlaMetArgArgLeuSerLysLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
Db 396 CCCGCGATCGCGCGCTCTCTATCTCTCAACAGAGAGCTTTTATCTGTGTTCAAGCTTA 455  
QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluValLys 120  
Db 456 GACCAACCGGACTCATTCGAGAGGTGCAAGAGCTCAACACAGATCTTACAGACCAAG 515  
QY 121 SerGlyLeuLysAsnLysThrLysGluAlaAlaGluLeuProMetValIleCysGlyAsn 140  
Db 516 TCCCTCTTCACAAACAAACAAAGAGATGTGAGTCCGCTGTCATTGCGGTAC 575  
QY 141 LysAsnAspHisGlyLeuLeuCysArgGlnValProThrThrGluAlaGluLeuVal 160

Db 576 AAAGGAGACCGG---GACTTCTACCGGGAAGTAGACAGCGGAGATGACGCTGTG 632  
QY 161 SerGlyAsp---GluAsnSerAlaTyrPheGluValSerLysLysValAsnThrAsnVal 179  
Db 633 GGTGACACACCTTCAGCCCTTGTGCTTCTGAGATCTCAGACCAAGAAAGACAGCTTG 692  
QY 180 AspGluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAla 199  
Db 693 GACCAAGATCTTCCTGCGGCTCTTTGCCATGCGCCAGCTGCTTACAGAGATGAGCCCGAC 752  
QY 200 LeuHisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMet 219  
Db 753 TTGCACCGCAGAGATGTCGTGACACTGACGACCTTACTGACAGAAAG-----GCTCTG 806  
QY 220 ArgArgValLysGluMet-----AspAlaTyr 228  
Db 807 AGGAACAGAAAGCTTCTGCGTGGCGGACGCGAGCGCGGCGGACCAACGCGGATGCTT 866  
QY 229 GlyMetValSerProPheAlaArgArgProSerValAsnSerAspLeuLysTyrIleLys 248  
Db 867 GGCATCTTGCGGCGCTTGTCTCGAGACCCAGCTTGCACGACCTCATATGATTCATTCGT 926  
QY 249 AlaLysValLeuArgGluGlyGlnAlaArgGluArgLysLysCysThrIle 265  
Db 927 GAAAAAACCACTGTGCGGACGAGCTAAGCAAGAGCGCTGTCTATC 977  
RESULT 11  
CF999619  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLES  
JOURNAL  
COMMENT  
CF999619 761 bp mRNA linear EST 25-NOV-2003  
AGNCOURT.1631970 NIH\_ZGC\_7 Danio rerio cDNA clone IMAGE:7039679  
5', mRNA sequence.  
CF999619  
CF999619.1 GI:38520470  
EST.  
Danio rerio (zebrafish)  
SOURCE  
Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 761)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [csgerba@mail.nih.gov](mailto:csgerba@mail.nih.gov)  
Tissue Procurement: Jen Zou, Harvard  
cDNA Library Preparation: Open BioSystems  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: ILAM14793 row: e column: 21  
High quality sequence stop: 687.  
Location/Qualifiers  
1..761  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:7039679"  
/tissue\_type="whole body"  
/lab\_host="DH10B"  
/note="Vector: pEXpress1, Site 1: NotI; Site 2: EcoRV;  
Bulk tissue was collected from a whole adult individual  
from the Tuebingen strain. 1st strand cDNA was primed with  
a Not I - oligo(dT) primer, double-stranded cDNA was  
cloned into the Not I and EcoRV sites of pEXpress-1.  
library was size-selected for >1 kb fragments and  
normalized. A non-normalized version of this library is

## FEATURES

## SOURCE

also available (NIH ZGC 10). Library was constructed by  
Open Biosystems (Huntsville, AL)"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,05e-101 Length: 761  
Score: 869.50 Matches: 169  
Percent Similarity: 84.02% Conservative: 36  
Best Local Similarity: 69.26% Mismatches: 38  
Query Match: 63.14% Indels: 2  
Gaps: 1

US-09-778-963A-2 (1-266) x CF999619 (1-761)

```

Qy 1 MetMetlysthrleuSerSerGlyAsnCystrleuSerValProAlaIysAsnSerTyr 20
Db 32 ATGATCAAGAAAAGAGATCCGTCGGAACAGATTGACATCCCGCCAAAACCTGATC 91
Qy 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerArgPheIeu 40
Db 92 CGAGTGTGATCTCTCGGTCACCAAGTCCGGAGAGCGCGATCTGTCCGCTTTCTG 151
Qy 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHsaArgIysValTyr 60
Db 152 AACGAGCGCTTTGAGAGACATGACCGCCGACCATTTAGAGATTGATAGAACTTAC 211
Qy 61 AsnIleArgGlyAspMetTyrGlnIleuAspIleuAspThrSerGlyAsnHsaProPhe 80
Db 212 AGCATTAAGAGAGATGTGTATCAGCTGATTTTGTGATCTTACGAGAACCATTCATTC 271
Qy 81 ProAlaMetArgValGluSerSerIleleuThrGlyAspValPheIleleuValPheSerIeu 100
Db 272 CCAGCTATGCGGAGACATCTCATTTCTGACCGGTATGATGTTTATCTCTGTTTCACTTG 331
Qy 101 AspAsnArgGluSerPheAspGluValIysArgIleuGlnIysGlnIleuGluValIys 120
Db 332 GACAAACCGCGAGTCTTCCAGAGTGACGCTGAGAGACAGATCTTACGAGACCAA 391
Qy 121 SerCysleuIysAsnIysThrIlyGluAlaIleGluIeuProMetValIleCysGlyAsn 140
Db 392 TCCGCTTTAAACCAAGACCAAGAACGTCGACGCTGCTGTGATCTGCGGGGAC 451
Qy 141 LysAsnAspHisGlyIleuGlnIysArgIleuValProThrThrIleGluAlaIleuIleuVal 160
Db 452 AAGGCGCAACCC---GAGTTTACCGAGAGATTACGCCCGACGAGATGACAGCTGATC 508
Qy 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIysIysAsnThrAsnValAsp 180
Db 509 GCCGAGACGAGCAATGGCGTACTTGAAGATCTCCGCTTAAAGAAACAGACATGAT 568
Qy 181 GluMetPheTyrValIleuPheSerMetAlaIysIeuProHsaGluMetSerProAlaIeu 200
Db 569 CAGATGTTCAGACACTTTTACGCTCCGCAAACTCCCAATGAGATGAGCCCTGATCTC 628
Qy 201 HisArgIysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220
Db 629 CATGCAAGAGTGTCGAGTCACTGATCTGATCTGCAATTAAGTCCCTGAA-GAACAA 687
Qy 221 ArgValIysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240
Db 688 GAATGAAAGACGCGGCGTACGCGATGTCCGCGCTTTCCGCGCGCGCCGAGCGTG 747
Qy 241 AsnSerAspIeu 244
Db 748 CACAGTACTTA 759

```

## RESULT 12

LOCUS BX327655 891 bp mRNA linear EST 01-MAY-2003  
DEFINITION BX327655 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0D0015YK10 5-PRIME, mRNA sequence.  
ACCESSION BX327655  
VERSION BX327655.1 GI:30307667

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS 1 (bases 1 to 891)  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
JOURNAL Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10498.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BA0072D04\_C  
S00604.1cluster=10498.f. Contact : Feng Liang Email :  
liang@lifetech.com URL : http://fulllength.invitrogen.com/  
Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID :  
CS0BA0072D04\_CS00604.1.

## FEATURES

## source

Location/Qualifiers

1..891

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D0015YK10"

/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="First strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

## Alignment Scores:

Pred. No.: 3.1e-100 Length: 891  
Score: 859.00 Matches: 179  
Percent Similarity: 91.92% Conservative: 3  
Best Local Similarity: 90.40% Mismatches: 12  
Query Match: 62.38% Indels: 4  
Gaps: 2

US-09-778-963A-2 (1-266) x BX327655 (1-891)

```

Qy 1 MetMetlysthrleuSerSerGlyAsnCystrleuSerValProAlaIysAsnSerTyr 20
Db 306 ATGATGAAGACTTTGTCCAGCGGAGACTGACGCTGATGCGCCGCAAAAACATCATAC 365
Qy 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerArgPheIeu 40
Db 366 CGCATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 425
Qy 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHsaArgIysValTyr 60
Db 426 AATGCGCGCTTTGAGAGACATGACACCACTTCAAGAGCTTCCACCGTAAGTATAC 485
Qy 61 AsnIleArgGlyAspMetTyrGlnIleuAspIleuAspThrSerGlyAsnHsaProPhe 80
Db 486 AAGATCGCGCGGAGACATGATACCACTCGATCTGATCTTGGGACCAACACCCCTTC 545
Qy 81 ProAlaMetArgValGluSerSerIleleuThrGlyAspValPheIleleuValPheSerIeu 100
Db 546 CCCGCAATGCGCAGAGCTGTCCATCTTCAAGAGAGATGCTTTCATCTGTGTTCAAGCTTG 605
Qy 101 AspAsnArgGluSerPheAspGluValIysArgIleuGlnIysGlnIleuGluValIys 120
Db 606 GATAACCGGAGATCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 665
Qy 121 SerCysleuIysAsnIysThrIlyGluAlaIleGluIeuProMetValIleCysGlyAsn 140
Db 666 TCTGCTGAAGAGACCAAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 725

```

Qy 140 nlyasnaaphisgllyuleuCyarglvalProthrThrglualagluleuVa 160  
Db 726 CAAGAAACGACCGAGCTGTGCGCCAGGTGCCACCGAGGCGGAGCTGTGT 785  
Qy 160 lserglvalspgluanserialATyPhglvalserlalylysAsnVala 180  
Db 786 GTCGGGCGACGAGGAGCTGCTTGTGAGGTGTGCGCCAAAGAG--ACACCACTGTG 842  
Qy 180 spglumetPheTyValleupheserMetAlalySleuproHigluMet 196  
Db 843 ACAGAGATTCTTACGTGCTTACGATG---CCAACGACACAGATG 889  
RESULT 13  
CF997752 726 bp mRNA linear EST 25-NOV-2003  
LOCUS CF997752  
DEFINITION AGENCOURT 16394218 NIH\_ZGC\_7 Danio rerio cDNA clone IMAGE:7039416  
5' mRNA sequence.  
ACCESSION CF997752.1 GI:38518603  
VERSION CF997752.1 GI:38518603  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 726)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: c9agbs-r@mail.nih.gov  
Tissue Procurement: Len Zon, Harvard  
cDNA Library Preparation: Open Biosystems  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ULNL at:  
http://image.llnl.gov  
Plate: L1M41792 row: 3 column: 22  
High quality sequence stop: 627.  
Location/Qualifiers  
FEATURES  
source  
1..726  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:7039416"  
/tissue\_type="whole body"  
/lab\_host="DH10B"  
/note="Vector: pEXpress1; Site 1: NotI, Site 2: EcoRV;  
Bulk tissue was collected from a whole adult individual  
from the Tuebingen strain, 1st strand cDNA was primed with  
a Not I - oligo(dT) primer, double-stranded cDNA was  
cloned into the Not I and EcoRV sites of pEXpress-1.  
Library was size-selected for >1 kb fragments and  
normalized. A non-normalized version of this library is  
also available (NIH ZGC 10). Library was constructed by  
Open Biosystems (Huntsville, AL)"

## ALIGNMENT SCORES:

Pred. No.: 1.17e-99 Length: 726  
Score: 853.50 Matches: 161  
Percent Similarity: 83.90% Conservative: 37  
Best Local Similarity: 68.22% Mismatches: 37  
Query Match: 61.98% Indels: 1  
DB: 14 Gaps: 1

US-09-778-963A-2 (1-266) x CF997752 (1-726)

Qy 1 MetwelythrleuSerSerGlyAenCyThrleuSerValProAlalyAsnSerTy 20  
Db 21 ATGATCAAGAAATGATCCGTGGAAACAGATTGACATCCCGGAGAACTGTAC 80  
Qy 21 ArgMetValleuGluAlaSerArgValGlyLysSerSerIleValSerArgPhe 40  
Db 81 CGATGTGATTCCTCGGTCCACCAAGTCGGGAAGACGCGATCGTGTGCGCTTCTG 140  
Qy 41 AsnGlyArgPheGluAspGlnTyThrProThrIleGluAspPheHisArgValTy 60  
Db 141 AACCGAGCTTGTGGAGCAGTACACCGCACCATTTGAGGACTTTCATGAGAACTTAC 200  
Qy 61 AsnIleArgGlyAspMetTyThrLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
Db 201 AGCATTTAAGCAGATGTGTATCAGCTGATTTTGTGATCTTCAGGAAACATTCATTC 260  
Qy 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
Db 261 CCAGCTATGCGGAGACTGTCCATTCGACCGGTATGTATTATCCTGTGTTCACTTTC 320  
Qy 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluValLys 120  
Db 321 GACACCGCGAGTCTTCCACGAGGTGACGGTCTGAAGCAGCATCTACAGACCAAAA 380  
Qy 121 SerCysLeuLysAsnLysThrLysGluAlaAlaGluLeuProMetValIleCysGlyAsn 140  
Db 381 TCTGCTTTAAACAAAGACCAAGAACACGATGACGTGCTTCGTGTATGTGCGGAGAC 440  
Qy 141 LysAsnAspHisgllyuleuCyarglvalProthrThrglualagluleuVal 160  
Db 441 AAGGCGACGCC--GAGTTTACCGAAGTTGACGCGCAGACGACGACGACCTATC 497  
Qy 161 SerGlyAspGluAsnSerAlaTyPheGluValSerAlalyLysAsnThrAsnValAsp 180  
Db 498 GCCGAGACGAGCAATGCGCTACTGAGATCTCCGCTTAAAGAAACACACCAAGCTGAT 557  
Qy 181 GluMetPheTyValleupheserMetAlalySleuproHigluMetSerProAlaLeu 200  
Db 558 CAGATGTTCCAGAGACTTTTACCTGCGCAAACTGCCAATGATGAGCCCTGATCTC 617  
Qy 201 HisArgLysIleSerValGlnTyGlyAspAlaPheHisProArgProPheCysMetArg 220  
Db 618 CATCGNCAAGTGTCCGTGACTGTCATCTCTGCAATMAAAGTCGTCGAGAACAG 677  
Qy 221 ArgValLysGluMetAspAlaTyThrGlyMetValSerProPheAlaArg 236  
Db 678 AAGGTGAAGACGCGGACGCGATGCGATGTGCGGCTTTCGCGCGC 725  
RESULT 14  
CD513046 834 bp mRNA linear EST 06-JUN-2003  
LOCUS CD513046  
DEFINITION AGENCOURT 14371657 NIH\_MGC\_179 Homo sapiens cDNA clone  
IMAGE:30394422 5' mRNA sequence.  
ACCESSION CD513046  
VERSION CD513046.1 GI:31444764  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 834)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: c9agbs-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: Agencourt Bioscience Corporation



US-09-778-963A-2 (1-266) x CK029404 (1-749)

```
QY      1 MetMetIySthrlEuSerSergIyAsnCysThrlEuSerValProAlaIyAsnSerTyr 20
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      48 ATGATCAAGAAATAGTCCGTCGGAAAAACGAGTTTGAATCCCGCGGAAAAAAGCTGTAC 107
QY      21 ArgMetValIalLeuGIyAlaSerArgValGIyIySerSerIleValSerArgPheLeu 40
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      108 CGATGTGATCTCTCGGGTCCACCAAGTCGGGAAGACGGCGATGTGTGGGCTTTCTG 167
QY      41 AsnGIyArGPheGIyAspGIyAlnTyrThrProThrlIleGIyAspPheIleArgIyValTyr 60
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      168 AACGGACCTTTGAGAGAGATACCGCCGACATTGAGACTTTCATAGGAAACTTAC 227
QY      61 AsnIleArGIyAspMetTyrGIyAlnLeuAspIleLeuAspThrSergIyAsnIleProPhe 80
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      228 AGCATTAAAGGATGTGTATCAGCTGGATATTTGGATATTTCAGGGAAACCATCATTC 287
QY      81 ProAlaMetArgArgLeuSerIleLeuThrgIyAspValPheIleLeuValPheSerLeu 100
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      288 CCAGCTATGCCGAGACTGTCCATTCGACCCGATGTGTATCTCGGTGTTCAGTTG 347
QY      101 AspAsnArgGIyLeuSergPheAspGIyValIyAsArgLeuGIyAlnIleuGIyValIyS 120
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      348 GACAAACCGGAGTCTTTCACAGAGGTGCAGCTCTGAAAGACAGATCTACAGAGACAA 407
QY      121 SerCysLeuIyAsnIySthrlEuSgluaIalagIuLeuPrometValIleCysGIyAsn 140
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      408 TCCTGCCCTTAAACAAAGCAAGAAAGAAAGTGCAGTGCCTGTGTGTCTGCGGGAAC 467
QY      141 LysAsnAspHisGIyGIuLeuCyArgIyAlnValProThrlThrgIuAlagIuLeuVal 160
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      468 AAGGGCGACCGC--GAGTTTACCGAGAAAGTTCAAGCGGACGAGATCGAGCACTGATC 524
QY      161 SerGIyAspGIyAsnSerAlaTyrPheGIyValSerAlaIySgluIyAsnThrAsnValAsp 180
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      525 GCCGAGAGACGACATGCGCTACTTCGAGATCTCCGCTAAAGAGACGAAACGTGAT 584
QY      181 GIuMetPheTyrValLeuPheSerMetAlaIySleuProHisGIuMetSerProAlaLeu 200
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      585 CAGATGTTCAGAGACTCTTACGCTCGCCAAACTGCCCAATGAGATGAGCCCTGATCTC 644
QY      201 HisArgIyIleSerValGIyTyrGIyAspAlaPheHisProArgProPheCysMetArg 220
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      645 CATCGAAGGTGTCCGTGCACTGTGTGACATCTGCATTA--AAGTCCTGAAGAACAAAG 703
QY      221 ArgValIySgluMetAspAlaTyrGIyMetValSerPro 233
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      704 AAGGTGAAGACGGC-GACCGCTACGGCATTTGTGCGGCT 741
```

Search completed: June 20, 2004, 22:57:58

Job time : 2780 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 20, 2004, 17:41:52 ; Search time 4194 Seconds

(without alignments)  
2748.986 Million cell updates/sec

Title: US-09-778-963A-2

Sequence: 1 MMKTLSGNCNTLSVPAKNSY.....IKAKVLEGGQARDKCTIQ 266

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgm2\_1/USPTO.spool/US09778963/runat\_18062004\_123459\_15897/app\_query.fasta\_1.455  
-DB=GenBdb1 -OFMT=faaftp -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09778963.QCGN\_1\_13731.@runat\_18062004\_123459\_15897 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGOUDRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBdb1.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vl.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrc.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1377	100.0	3082	6	AX572848 Sequence
2	1372	99.6	2832	6	AX393244 Sequence
3	1372	99.6	2832	6	AF279143 Homo sapi
4	1372	99.6	2973	6	AX393267 Sequence
5	1372	99.6	3049	6	BC013419 Homo sapi
6	1372	99.6	3427	6	AX721069 Sequence
7	1304	94.7	2824	10	BC036988 Mus muscu
8	1304	94.7	3020	6	AX393362 Sequence
9	1304	94.7	3469	10	AF134409 Rattus no
10	1201	87.2	2699	9	HSB803172 Homo sapi
11	955.5	69.4	210264	2	AC145933 Gallus ga
12	942.5	68.4	11390	5	BC056272 Danio rer
13	911	66.2	11221	6	AX572850 Sequence
14	911	66.2	114771	9	HS569D19 Human DNA
15	906.5	65.8	178071	2	AC145536 Lemur cat
16	903	65.6	274402	2	AC145758 Microcebu
17	879	63.8	846	9	AF498923 Homo sapi
18	879	63.8	979	9	AF172846 Homo sapi
19	879	63.8	1187	9	AF153192 Homo sapi
20	879	63.8	1740	6	AF780438 Sequence
21	879	63.8	1740	9	AF069506 Homo sapi
22	879	63.8	1744	6	BD190304 Elx1 phos
23	879	63.8	1758	9	BC018041 Homo sapi
24	879	63.8	1841	6	AR236058 Sequence
25	876.5	63.7	1616	10	AF239157 Rattus no
26	876.5	63.7	1689	6	AR236060 Sequence
27	874.5	63.5	1612	10	BC034166 Mus muscu
28	874.5	63.5	1623	10	AF009246 Mus muscu
29	871	62.9	1776	6	BD248751 GTPase as
30	866	62.9	1746	9	AF177335 Homo sapi
31	855	62.1	211075	10	AC076974 Mus muscu
32	853	61.9	286789	2	AC133723 Rattus no
33	839	60.9	179124	10	AL603710 Mouse DNA
34	839	60.9	247899	2	AC025909 Mus muscu
35	820	59.5	3079	6	AR236061 Sequence
36	820	59.5	314307	2	AC122995 Rattus no
37	817	59.3	3986	6	AR236059 Sequence
38	817	59.3	4990	9	AF222979 Homo sapi
39	817	59.3	5141	9	AF262018 Homo sapi
40	817	59.3	141048	9	AC073621 Homo sapi
41	817	59.3	183334	9	AC020558 Homo sapi
42	817	59.3	183427	2	AC090608 Homo sapi
43	714	51.9	217841	2	AC095995 Rattus no
44	686	49.8	260894	2	AC119873 Mus muscu
45	676	49.1	2294	10	BC026377 Mus muscu

RESULT 1

## ALIGNMENTS

AX572848  
 LOCUS AX572848 3082 bp DNA linear PAT 29-NOV-2002  
 DEFINITION Sequence 1 from Patent WO02062849.  
 ACCSSION AX572848  
 VERSION AX572848.1 GI:26004937  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Neelan, B., Ketchum, K.A., di Francesco, V. and Beasley, E.M.  
 AUTHORS Isolated human ras-like proteins, nucleic acid molecules encoding  
 TITLE them, and uses thereof  
 JOURNAL Patent: WO 02062849-A 1 15-AUG-2002;  
 PE Corporation (NY) (US)  
 FEATURES  
 source location/Qualifiers  
 1..3082  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores: 3.03e-126 Length: 3082  
 Pred. No.: 1377.00 Matches: 266  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 6 Gaps: 0  
 DB: 6  
 US-09-778-963a-2 (1-266) x AX572848 (1-3082)  
 QY 1 MetMettlysrthleuSerSerGlyAsnCystrhLeuSerValProAlaIlyAsnSerTyr 20  
 Db 211 ATGATGAGAGACTTTGTCACGCGGAACTGCACAGCTCAGTGTGCCGCCAAAACCTCATAC 270  
 QY 21 ArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40  
 Db 271 CGCATGTGTGTCTGGGTCTCTCGGGTGGGAGAGCTCCATCTGTGTCTGCTTCTC 330  
 QY 41 AengLYArgPheGlyuAspGlnTyrThrProThrIleGlyuAspPheHisArgLYValTyr 60  
 Db 331 AATGGCCGCTTGAGAGACAGTACACACCCACATGAGAGACTTCCACCGTAAGATAC 390  
 QY 61 AenlleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 Db 391 AACATCCGCGGACATGTACACAGCTGCATCTCGATACCTCTGGCAACACCCCTTC 450  
 QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 Db 451 CCGGCATGTGGAGGCTGTTCATCTCCACAGGGAGATGTCTTCACTCTGTGTTCAGCTG 510  
 QY 101 AspAsnArgGlyuSerPheAspGlyValIlyArgLeuGlnLysGlnIleLeuGlyValIly 120  
 Db 511 GATTAACGGGAGATCTTCATGATGAGTCAAGCGCTTCAGAAAGACATCTCGAGAGTCAAG 570  
 QY 121 SerCysLeuLysAsnIlysrthlysrGlyuAlaIleGlyLeuProMetValIleCysGlyAsn 140  
 Db 571 TCTGTCTGTGAAGAACAGACAGAGAGGCGCGAGCTGCCATGTGCTGTGTGCAAC 630  
 QY 141 LysAsnAspHisGlyGlyLeuGlyCysArgGlnValProThrThrGlyuAlaGlyLeuLeuVal 160  
 Db 631 AAGAAAGACACACCGGACAGCTGTGTGCCCGCAGTCCACACGAGGCGGAGCTGGTGG 690  
 QY 161 SerGlyAspGlyuAsnSerAlaTyrPheGlyuAlaSerAlaLysIlysrthlysrAsnValAsp 180  
 Db 691 TCGGCGAGACAGAACTCCGCTTCTCGAGGTGTCCGCAAGAAAGACCAACAGTGGAC 750  
 QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGlyuMetSerProAlaLeu 200  
 Db 751 GAGATGTTTACGTGTCTTCAGATGTGCAAGCTGCCACACAGATGAGCCCGCCCTG 810

QY 201 HisArgHisIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
 Db 811 CATGCAAGATCTTCGTCAGTACGAGTACGAGCTTCCACCCAGGCGCTTGTGCATGCGC 870  
 QY 221 ArgValIlysrGlyuMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240  
 Db 871 CGGTCAGAGAGATGACAGCGCTTATGCGATGTCTCGCCCTTCCGCCGCCCGCCAGCGTC 930  
 QY 241 AsnSerAspLeuIlysrTyrIleValAlaIlyValLeuArgGlyuGlnAlaArgGlyuArg 260  
 Db 931 AACAGTACCTCAGATCATCAAGAGCCAGAGTCTTCGGAAGGCCAGGCGCTGTAGAGG 990  
 QY 261 AspLysCysThrIleGln 266  
 Db 991 GACAAATGCACCATTCAG 1008  
 RESULT 2  
 AX393244 2832 bp DNA linear PAT 23-MAR-2002  
 LOCUS AX393244  
 DEFINITION Sequence 174 from Patent WO0210217.  
 ACCSSION AX393244  
 VERSION AX393244.1 GI:19701294  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 St Croix, B., Kinzler, K.W. and Vogelstein, B.  
 AUTHORS Endothelial cell expression patterns  
 TITLE Patent: WO 0210217-A 174 07-FEB-2002;  
 JOURNAL The Johns Hopkins University (US)  
 FEATURES  
 source location/Qualifiers  
 1..2832  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores: 8.54e-126 Length: 2832  
 Pred. No.: 1372.00 Matches: 265  
 Score: 99.62% Conservative: 0  
 Percent Similarity: 99.62% Mismatches: 1  
 Best Local Similarity: 99.64% Indels: 0  
 Query Match: 6 Gaps: 0  
 DB: 6  
 US-09-778-963a-2 (1-266) x AX393244 (1-2832)  
 QY 1 MetMettlysrthleuSerSerGlyAsnCystrhLeuSerValProAlaIlyAsnSerTyr 20  
 Db 37 ATGATGAGAGACTTTGTCACGCGGAACTGCACAGCTCAGTGTGCCGCCAAAACCTCATAC 96  
 QY 21 ArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40  
 Db 97 CGCATGTGTGTCTGGGTCTCTCGGGTGGGAGAGCTCCATCTGTGTCTGCTTCTC 156  
 QY 41 AengLYArgPheGlyuAspGlnTyrThrProThrIleGlyuAspPheHisArgLYValTyr 60  
 Db 157 AATGGCCGCTTGAGAGACAGTACACACCCACATGAGAGACTTCCACCGTAAGATAC 216  
 QY 61 AenlleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
 Db 217 AACATCCGCGGACATGTACACAGCTGCACATCTCGATACCTCTGGCAACACCCCTTC 276  
 QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100  
 Db 277 CCGGCATGTGGAGGCTGTTCATCTCCACAGGGAGATGTCTTCACTCTGTGTTCAGCTG 336  
 QY 101 AspAsnArgGlyuSerPheAspGlyValIlyArgLeuGlnLysGlnIleLeuGlyValIly 120  
 Db 337 GATTAACGGGAGATCTTCATGATGAGTCAAGCGCTTCCAGAAAGACATCTCGAGAGTCAAG 396

Alignment Scores:  
 Pred. No.: 8,54e-126 Length: 2832  
 Score: 1372.00 Matches: 265  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.64% Indels: 0  
 Db: 9 Gaps: 0

US-09-778-963A-2 (1-266) x AF279143 (1-2832)

QY	1	MetMetLysThrLysSerSerGlyYAsnCySerThrLysSerValProAlaLysAsnSerTyr	20
Db	37	ATGATGAAGACTTTGTCTCAGCGGGAACTGCAGCGCTCAGTGTGCGCCGCAAAACCTCATAC	96
QY	21	ArgMetValValLeuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu	40
Db	97	CGCATGGTGGTGCTGGGTGCTCTCGGGTGGGCAAGAGCTCATCGTGTCTCGTTCTTC	156
QY	41	AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgGlyValTyr	60
Db	157	AATGGCCGCTTTGAGGACCAAGTACACACCACCATCCAGGACTTCCACCGTAAAGTATAC	216
QY	61	AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe	80
Db	217	AACATCCGGCGGCAATGTACCACTCGACATCTCGATATCTTGAGCAACACCCCTTC	276
QY	81	ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu	100
Db	277	CCCGCGATCGCCCAAGCTGTCTCATCTCTCAGAGGGATGTCTTCATCTGGTGTTCAGCTG	336
QY	101	AspAsnArgGlnSerPheAspGlnValIleArgLeuGlnLysGlnIleLeuGluValLys	120
Db	337	GATTAACCGGGAGTCTCTTCATGAGAGTCAAGCGCTTCAGAAAGCATCTTGAGAGTCAAG	396
QY	121	SerCyLeuLysAsnLysThrLysGluAlaAlaGluLeuProMetValIleCyGlyYAsn	140
Db	397	TCTCGGCTCGAAGAACAGACCAAGAGAGCGCGGAGCTGCCATGTGTATGTGTGGCAAC	456
QY	141	LysAsnAspHisGlyGlnLeuCyArgGlnValProThrThrGlnAlaGlnLeuVal	160
Db	457	AAGAACAACACCGCGAGCTGTGCCCGCAGGTGCCACACCGAGCGCGAGCTCTGGTG	516
QY	161	SerGlyAspGluAsnSerIleTyrPheGluValSerAlaLysIleYAsnThrAsnValAsp	180
Db	517	TCGGGCGACGAGAACTGGGCTCTTCAGAGGTGTGGGCAAGAGAACACCAACGTTGAC	576
QY	181	GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu	200
Db	577	GAGATGTTCTACGGTCTCTTCAGATAGGCCAAAGTGCACACGAGATGAGGCCGCCCTG	636
QY	201	HisArgLysIleSerValGlnTyrGlyYAspAlaPheHisProArgProPheCyMetArg	220
Db	637	CATGCGAAGATCTCTCGTGCAGTACGGGTGAGCGCTTCCACCCAGGCCCTTGTGATGGC	696
QY	221	ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal	240
Db	697	CGCGTCAAGAGAGTAGACGCTTATGGCATGTGTCTGCCCTTCCGCCCGCCAGCGCTC	756
QY	241	AsnSerAspLeuLysTyrIleLysAlaLysValLeuArgGlnLysGlnAlaArgGluArg	260
Db	757	AACAGTAGCTCCAAGTACATCAAGGCCAAGGTCTCTTCGGGAAGGCCAAGGCCCGTGAAGG	816
QY	261	AspLysCyThrIleGln	266
Db	817	GACAAAGTGCACATTCAG	834

```

RESULT 4
AX393267          2973 bp   DNA      linear      PART 23-MAR-2002
LOCUS             Sequence 197 from Patent WO0210217.
DEFINITION        AX393267
ACCESSION         AX393267
VERSION           AX393267.1  GI:19701305
KEYWORDS
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 St Croix,B., Kinzler,K.W. and Vogelstein,B.
  Endothelial cell expression patterns
  Patent: WO 0210217-A 197 07-FEB-2002;
  The Johns Hopkins University (US)
FEATURES
  source
    1..2973
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
  Pred. No.:      9, 07e-126      Length:      2973
  Score:          1372.00         Matches:      265
  Percent Similarity: 99.62%      Conservative: 0
  Best Local Similarity: 99.62%   Mismatches:  1
  Query Match:    99.64%         Indels:       0
  Ds:              6              Gaps:         0
US-09-778-963A-2 (1-266) x AX393267 (1-2973)
QY      1 MetMetLysThrLeuSerSerGlyAsnGlyThrLeuSerValProAlaLysAsnSerTyr 20
      208 ATGATGAAGACCTTTGTCACGGGAGACTCAGCTGATGCTGCGCCAAAACCTCATAC 267
QY      21 ArgMetValIleuLeuGlyIleAsnArgValGlyLysSerSerIleValSerArgPheLeu 40
      268 CCGCATGTGTGCTGTGGTCTCTCTGCGGTGGGCAAGAGCTCCATGCTGTCTGCTTCC 327
QY      41 AangGlyArgPheGluAsnGlnTyrThrProThrIleGluAsnPheHisArgLysValTyr 60
      328 AATGGCCGCTTGGAGGACGATGACACACACCATCGAGACTTCCACCGTAAGGTATAC 387
QY      61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
      388 AACATCCGGGGGCGACATGTAACAGCTCCGACATCTGGATACCTCTGGCAACACCCCTTC 447
QY      81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
      448 CCGGCATGCGGAGGCTGTCATCTCCACAGGGGATGCTTCATCCTGCTGTTACGCTTC 507
QY      101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluValLys 120
      508 GATTAACCGGAGACTCTTCATGATGAGGTCAAGCCCTTCAGAAACAGATCTGAGAGGTCAAG 567
QY      121 SerCysLeuLysAsnLysThrLysGlnLysIleValIleGluLeuProMetValIleCysGlyAsn 140
      568 TCTTCGCTGAAAGAACAGACAGACAGAGAGGCGGAGCTCCCATGCTATCTGTGGCAAC 627
QY      141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGluIleValLeuVal 160
      628 AAGAAGCAGACACAGCGGAGGCTGCGCGCAGGTGCCACACCGAGGCGAGCTGCTGGTG 687
QY      161 SerGlyAspGluAsnSerLysThrPheGluValSerAlaLysLysAsnThrAsnValAsp 180
      688 TCGGGCGAGAGAACTGCGCTTCACTTCGAGGTGTGGGCCAAGAAACACCAACGTGAGAC 747
QY      181 GluMetPheTyrValLeuPheSerMetAlaLysLysProHisGluMetSerProAlaLeu 200
      748 GAGATGTTCTACGTCTCTTTCAGCATGGCCAGAGCTGCCACAGATGAGCCCGCCCTTG 807

```

```

QY      201 HisArgLysIleSerValGlnTyrGlyAsnAlaPheHisProArgProPheCysMetArg 220
      808 CATCCAGAGATCTCCGTGACAGTACGTACCGCTTCCACCCAGGCCCTTCTGCAATGCGC 867
QY      221 ArgValIlyGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240
      868 CCGCTCAAGAGATGAGACGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
QY      241 AsnSerAspLeuLysTyrIleValIleValIleValIleValIleValIleValIleVal 260
      928 AACAGTACCTCAAGTAAATCAAGGCGCAAGGCTCTTGGAGAGGCGAGGCCCGCTAAGG 987
QY      261 AspLysCysThrIleGln 266
      988 GACAGTGCACCATCTCAG 1005
Db

```

```

RESULT 5
BC013419          3049 bp   mRNA      linear      PRI 03-OCT-2003
LOCUS             Homo sapiens RASD family, member 2, mRNA (cDNA clone MGC:4834
DEFINITION        IMAGE:3604196), complete cds.
ACCESSION         BC013419
VERSION           BC013419.2  GI:33870961
KEYWORDS
SOURCE            MGC.
ORGANISM          Homo sapiens (human)
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 3049)
  Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
  Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
  Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
  Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
  Ditchenko,L., Marusik,K., Farmer,A.A., Rubin,G.M., Hong,L.,
  Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
  Scheetz,T.E., Brownstein,M.J., Umed,T.B., Toshiyuki,S.,
  Carninci,P., Prange,C., Raha,S.S., Loguercio,N.A., Peters,G.J.,
  Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
  McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
  Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
  Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
  Fahey,J., Helton,E., Kettelman,W., Madan,A., Rodriguez,S.,
  Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
  Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
  Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
  Butlerfield,Y.S., Krzywinski,M.I., Skalska,J., Smailus,D.E.,
  Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.,
  Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences
  Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  22388257
  12477932
  2 (bases 1 to 3049)
  Strausberg,R.
  Direct Submission
  Submitted (31-AUG-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

REMARK
COMMENT
  NIH-MGC Project URL: http://mgc.nci.nih.gov
  On Aug 19, 2003 this sequence version replaced gi:15426590.
  Contact: MGC help desk
  Email: cga@pds-remail.nih.gov
  Tissue Procurement: ATCC
  CDNA Library Preparation: Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
  DNA Sequencing by: Institute for Systems Biology
  http://www.systemsbiology.org
  contact: amadan@systemsbiology.org
  Amp Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
  Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

```

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/INLNL at: <http://image.inl.gov>  
Series: IRAL Plate: 11 Row: b Column: 18  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 22027485.  
Location/Qualifiers

## FEATURES

## SOURCE

1. 3049

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:4834 IMAGE:3604196"  
/issue\_type="Uterus, endometrium adenocarcinoma"  
/clone\_id="NH MGC\_44"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"

## gene

1. 3049

/gene="RASD2"  
/note="synonyms: TEM2, MGC:4834, Rhes"  
/db\_xref="LocusID:23551"  
208. 1008  
/codon\_start=1  
/product="RASD family, member 2"  
/protein\_id="AAH13419.1"  
/db\_xref="GI:15426591"  
/db\_xref="LocusID:23551"

## CDS

/translation="MMKTLSSGNCCTLSYPAKNSYRMVYLGAARYKSIYVRPLNGRE  
EDQYPTIEDHRRVYVNRGMVOLDIDTSGNHPMPMRRLSLTGDFVFLPSLND  
RSDFDEYKRLQKULEYKSCIKRKTKEAEIPMTYCGKNDHGEICROVPTTEABLY  
SGDENCAIYFEVSAKNTNIDEMFYLVLSMALPHMSPALRKRKISVOYGDAPHRPFC  
MRYKENDAYGMVSPFARPSNSDLKIKAKVLRREGARBERDKCTIO"  
259. 780  
/note="RAS; Region: Ras subfamily of RAS small GTPases"  
/db\_xref="CDD:smart00173"

## misc\_feature

/note="RAS; Region: Ras subfamily of RAS small GTPases"  
/db\_xref="CDD:smart00173"

## ORIGIN

## Alignment Scores:

Pred. No.: 9.35e-126 Length: 3049  
Score: 1372.00 Matches: 265  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.64% Indels: 0  
DB: 9 Gaps: 0

US-09-778-963A-2 (1-266) x BC013419 (1-3049)

QY 1 MetMetLyThrLeuSerSerGlyAsnCySthrLeuSerValProAlaIysAsnSerTyr 20  
DB 208 ATGATGAAGACTTTGTCTCCAGCGGGAACCTGCACTGCTGCTCCGCCCAAAACCTCATAC 267  
QY 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerArgPheIeu 40  
DB 268 CGCATGGTGTGCTGGGCTCTCGGGTGGGCAAGACTCCATCGTGTCTGCTTCTC 327  
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheIleArgIysValTyr 60  
DB 328 AATGGCGCTTTGAGGACCACTACACACCAACCACTGAGACTTCCACGTAAGTATAC 387  
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80  
DB 388 AACATCCCGCGCATGTACCACTGCACTCCGATACCTCTGGGCAACACCCCTTTC 447  
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerIeu 100  
DB 448 CCGCCCATGGCAGGCTCTCATCTCCACAGGGAGTGTCTTCACTTCTGTTCAAGCTTG 507  
QY 101 AspAsnArgGluSerPheAspGlnValIysArgLeuGlnTyrGlnIleLeuGlnValIys 120  
DB 508 GATTAACCGGAGTCTCTTGATGAGGTCAAGCGCTTCAAGAGCAATCTCGAGGTCAAG 567  
QY 121 SerCysLeuIysAsnIysThrIysGlnAlaIleGluLeuProMetValIleCysGlyAsn 140  
DB 568 TCCTGCCGTAAGAACCAAGCAAGGAGCGGCGGAGCTGCCCATGTGTATCTGTGGCAAC 627  
QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrThrGlnAlaGluLeuLeuVal 160

DB 628 AAGAACGACCAACGCGGAGCTGTGCGGAGGTGCCCAACCAAGCGCGAGCTGTGCTG 687  
QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaIysIysAsnThrAsnValAsp 180  
DB 688 TCGGCGCAGCAAGACTGTGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 747  
QY 181 GluMetPheTyrValIleuPheSerMetAlaIysLeuProHisGluMetSerProAlaIeu 200  
DB 748 GAGATGTTCTAGCGTCTCTTGAAGATGAGCTGCAAGCTGCAAGATGAGCTGCTGCTG 807  
QY 201 HisArgIysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220  
DB 808 CATGCAAGACTCTCGTGCATACGATGACGCTTCCACCCCAAGCCCTTGTGATGCGC 867  
QY 221 ArgValIysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240  
DB 868 CGCGTCAAGGAGATGAGCGCTTATGCAATGCTCTGCGCTTCCGCGCGCGCGCGCTTC 927  
QY 241 AsnSerAspLeuIysTyrIleValAlaIysValIleuArgGluGlnAlaArgGluArg 260  
DB 928 AACAGTGAAGCTTCAAGTACATCAAGGCAAGGTCTCTGAGAGGCGCAAGCGCGTGAAG 987  
QY 261 AspIysCysThrIleGln 266  
DB 988 GACAAGTGCACCATCCAG 1005

## RESULT 6

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## SOURCE

## ORIGIN

## Alignment Scores:

Pred. No.: 1.08e-125 Length: 3427  
Score: 1372.00 Matches: 265  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.64% Indels: 0  
DB: 6 Gaps: 0

US-09-778-963A-2 (1-266) x AX721069 (1-3427)

QY 1 MetMetLyThrLeuSerSerGlyAsnCySthrLeuSerValProAlaIysAsnSerTyr 20  
DB 644 ATGATGAAGACTTTGTCTCCAGCGGGAACCTGCACTGCTGCTGCTGCTGCTGCTG 703  
QY 21 ArgMetValIleuGlyAlaSerArgValGlyIysSerSerIleValSerArgPheIeu 40  
DB 704 CGCATGGTGTGCTGGGCTCTCGGGTGGGCAAGACTCCATCGTGTCTGCTTCTC 763

```

QY 41 AsnGlyArpHegluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60
Db 764 AATGGCGCGCTTGGAGGACAGTACACCCACCATCGAGACTTCCACCGTAAGGRTAC 823
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
Db 824 AACATCGCGCGCACATGATCGACAGCTCGACATCTGTGATCTCTGGCAACCAACCCCTTC 883
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerIleu 100
Db 884 CCGCGCATCGCGAGCGCTGTCTCATCTCCACAGGGAGATGTCTCATCTCGGTTCAGCGCTG 943
QY 101 AspAsnArgGlyIleSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluValLys 120
Db 944 GATTAACCGGGAGTCTTCGATGAGGTCAAGCGCTTCAGAAAGAGATCTCGAGGTCAAG 1003
QY 121 SerCysLeuLysAsnLysThrLysGluValAlaGluLeuProMetValIleCysGlyAsn 140
Db 1004 TCTTGCTGTGAAGAACAGACCAAGAGGCGCGAGCTGCCCATGTGTCACTGTGGCAAC 1063
QY 141 LysAsnAspHisGlyGluLeuCysArgGlnValProThrGlnLysGlnIleLeuVal 160
Db 1064 AATAAGACCAACCGCGCATGCTGTCTCCCGCAGTCCCAACCAAGCGCGAGCTCTGTGAGT 1123
QY 161 SerGlyAspGluAsnSerIleTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180
Db 1124 TCGGGCGAGAGAACTGCGGCTACTCTCGAGGTGTCCGCCAAGAAACCAACGTCGAGAC 1183
QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
Db 1184 GACATGTCTACGTCGTCTTCAGCAAGCGCAACACTCCCAACAGATGAGGCCCGGCCCGT 1243
QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220
Db 1244 CATCGCAAGATCTCCGTGAGTACGATGAGCGCTTCACCCAGCGCGCTTCGATGAGCGCC 1303
QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgProSerVal 240
Db 1304 CCGCTCAAGAGATGAGCGCGCTATGATGCTGTCTCCCGCTCGCGCGCGCGCGCGCGT 1363
QY 241 AsnSerAspLeuLysTyrIleLysAlaLysValLeuArgGluGlnAlaArgGluArg 260
Db 1364 AACAGGACCTCAAGTACATCAAGGCGCAAGTCTCTTGAGAGAGCCAGCGCGCTGAGAG 1423
QY 261 AspLysCysThrIleGln 266
Db 1424 GACAAGTGCACCATCCAG 1441

```

RESULT 7  
LOCUS BC036988 2824 bp mRNA linear ROD 16-APR-2003  
DEFINITION Mus musculus RIKEN cDNA 4930526B11 gene, mRNA (CDNA clone IMAGE:4988981), partial cds.  
ACCESSION BC036988  
VERSION BC036988.1 GI:22477555  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
1 (bases 1 to 2824)  
Strausberg, R.L., Feigold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altchul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Kana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Wallaby, S.J., Bosak, S.A., McEwan, P.J., McEwen, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

FEATURES  
source  
CDS  
gene  
CDS

REMARK  
COMMENT  
JOURNAL  
MEDLINE  
PUBMED  
2 (bases 1 to 2824)  
12477932  
12477932  
TITLE  
AUTHORS  
JOURNAL  
Submitted (23-AUG-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooke, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haplight, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, O.L., Mastello, C., Mackeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, D., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>  
Series: TRAK Plate: 81 Row: 9 Column: 15  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction.  
Location/Qualifiers  
1..2824  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4988981"  
/tissue\_type="Colon, normal, 5 month old male mouse."  
/clone\_lib="NCI CGAP\_C024"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..2824  
/gene="4930526B11R1k"  
/note="synonym: Rasq2"  
/db\_xref="locusid:75141"  
/db\_xref="WGI:1922391"  
1..956  
/gene="4930526B11R1k"  
/codon\_start=3  
/product="4930526B11R1k protein"  
/protein\_id="AAH36988.1"  
/db\_xref="GI:22477555"  
/db\_xref="LOCUSID:75141"  
/db\_xref="WGI:1922391"  
/transcript="TRP1AVTLEPPOTPGDSAVGAGSPHLQELPGSPSPRAGPS  
RCQPRAMKTLSSGCTTAVLPAKNSYRVLGASRVKSSIVSRFLNGRFEQOYPT  
IEDHRVRYVHIDPMVOLDITLDGSHNPAPMARLITLGDVFLIVPSLIDRSFPEV  
KRIQKOLEIKYKTKTAEALPVLICNKRHDSELCQVYAMERABLVSDENCA  
YFEVSARKNTNVNEMTVLFSMAKLPHNSPALHKKISVYGDAPHRPCMKRTYA

## ORIGIN

GAGWVSPFARRPSVNSDLKYIKAKVLEQGAERDKSIC"

## Alignment Scores:

Pred. No.:	4,58e-119	Length:	2824
Score:	1304.00	Matches:	252
Percent Similarity:	96.24%	Conservative:	4
Best Local Similarity:	94.74%	Mismatches:	10
Query Match:	94.70%	Indels:	0
DB:	10	Gaps:	0

US-09-778-963A-2 (1-266) x BC036988 (1-2824)

```
QY 1 MetMetLysThrLeuSerSerGlyAsnCyThrLeuSerValProAlaLysAsnSerTyr 20
Db 156 ATGATGAAGACCTTGTCCAGTGGAACTGCACACTCAATGTGCTGCTAAGAACTCTTAC 215
QY 21 ArgMetValIalleuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40
Db 216 CGCATGTGTGTCTGTGGTGGCTCCGAGTGGGCAAGAGCTCATTTGTCTCCGCTTCTC 275
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60
Db 276 AATGCCCGCTTGAAGACCAAGTACAGCCCACTATCGAGACTTTCATCGCAAGGTGTAC 335
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
Db 336 AACATCCAGGGGACATGTACCAAGCTGGATCTCTGGACACTCTGGCAACCACTTC 395
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
Db 396 CTTGCCATGCGCGGCTCTCCATCTCTCAAGAGATGTCTTATCTGTGTGTTCAGCTG 455
QY 101 AspAsnArgGluSerPheAspGluValLysArgLeuGlnLysGlnIleLeuGluValLys 120
Db 456 GATAGCCCGGAGTCTTGTGATGAGGTCAAGCCCTCCAGAAACAGATCTGGAGGTCAAG 515
QY 121 SerCyLeuLysAsnLysThrIleGluIleAlaGluLeuProMetValIleCySerGlyAsn 140
Db 516 TCTTCCCTGAAGATTAACCAAGAGGCAAGAGCTGCCCCATGATGTGTGTGTGTGGAAAC 575
QY 141 LysAsnAspHisGlyGluLeuCyArgGlnValProThrThrGluAlaGluLeuVal 160
Db 576 AAGATGACCAAGAGAGCTGTGCGGCAAGTCCCTGCCATGAGAGGTGTGTGTGTGTGT 635
QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180
Db 636 TCTGTGTGTAAGAACTGCGGCTTCACTTCAAGGTGTGACAGCAAGAAACACTAATGTGAAT 695
QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
Db 696 GAGATGTTCTATGTCTGTTCAGCAAGGCAAGCTGCCCATGAGATGAGAGCTTGCACCTG 755
QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCyMetArg 220
Db 756 CACCATAGATCTCGTGTGACAGTACGCGCATGTCTTTCACCCCGGCTTTCGATGTGTGT 815
QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240
Db 816 CGCACTAAGTTCGACAGTGCCTATGAGCATGTGTCAACCTTTGCCGAGCGCCCAAGTGTTC 875
QY 241 AsnSerAspLeuLysTyrIleValAlaLysValLeuArgGluGlnValAlaArgGluLys 260
Db 876 AACAGTGTCTTAATATACATCAAGGCAAGGTCTTACGGAGAGGCGCAGAGCGGAGAGAGG 935
QY 261 AspLysCyThrIleGln 266
Db 936 GACAAGGTAGCATCCAG 953
```

RESULT 8  
AX393362 3020 bp DNA linear PAT 23-MAR-2002  
LOCUS AX393362  
DEFINITION Sequence 292 from Patent WO0210217.

## ACCESSION

AX393362  
AX393362.1 GI:19701352

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## SOURCE

## ORIGIN

## Alignment Scores:

Pred. No.:	4.98e-119	Length:	3020
Score:	1304.00	Matches:	252
Percent Similarity:	96.24%	Conservative:	4
Best Local Similarity:	94.74%	Mismatches:	10
Query Match:	94.70%	Indels:	0
DB:	6	Gaps:	0

US-09-778-963A-2 (1-266) x AX393362 (1-3020)

```
QY 1 MetMetLysThrLeuSerSerGlyAsnCyThrLeuSerValProAlaLysAsnSerTyr 20
Db 376 ATGATGAAGACCTTGTCCAGTGGAACTGCACACTCAATGTGCTGCTAAGAACTCTTAC 435
QY 21 ArgMetValIalleuGlyAlaSerArgValGlyLysSerSerIleValSerArgPheLeu 40
Db 436 CGCATGTGTGTCTGTGGTGGCTCCGAGTGGGCAAGAGCTCATTTGTCTCCGCTTCTC 495
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgLysValTyr 60
Db 496 AATGCCCGCTTGAAGACCAAGTACAGCCCACTATCGAGACTTTCATCGCAAGGTGTAC 555
QY 121 SerCyLeuLysAsnLysThrIleGluIleAlaGluLeuProMetValIleCySerGlyAsn 140
Db 496 AATGCCCGCTTGAAGACCAAGTACAGCCCACTATCGAGACTTTCATCGCAAGGTGTAC 555
QY 141 LysAsnAspHisGlyGluLeuCyArgGlnValProThrThrGluAlaGluLeuVal 160
Db 556 AACATCCAGGGGACATGTACAGTGTATCTCTGACACTCTGGCAACCACTTC 615
QY 161 SerGlyAspGluAsnSerAlaTyrPheGluValSerAlaLysLysAsnThrAsnValAsp 180
Db 616 CTTGCCATGCGCGGCTCTCCATCTCTCAAGAGATGTCTTCAATCTGTGTGTGTGTGTGT 675
QY 181 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
Db 676 GATAGCCCGGAGTCTTGTGATGAGGTCAAGGCGCTCCAGAAACAGATCTGGAGGTCAAG 735
QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCyMetArg 220
Db 736 TCTTCCCTGAAGATTAACCAAGAGGCAAGAGCTGCCCATGATGTATCTGTGTGTGTGT 795
QY 221 ArgValLysGluMetAspAlaTyrGlyMetValSerProPheAlaArgArgProSerVal 240
Db 141 LysAsnAspHisGlyGluLeuCyArgGlnValProThrThrGluAlaGluLeuVal 160
QY 241 AsnSerAspLeuLysTyrIleValAlaLysValLeuArgGluGlnValAlaArgGluLys 260
Db 796 AAGATGACCAAGAGAGCTGTGCGGCAAGTCCCTGCCATGAGAGGTGTGTGTGTGTGT 855
QY 261 AspLysCyThrIleGln 266
Db 856 TCTGTGTGTAAGAACTGCGGCTTATTTCAAGGTGTGACCAAGAAAGAACTAATGTGAAC 915
QY 261 GluMetPheTyrValLeuPheSerMetAlaLysLeuProHisGluMetSerProAlaLeu 200
Db 916 GAGATGTTCTATGTCTGTTCAGCATGTGCAAGTGTGCGCCCATGAGATGAGCGCTGTGATG 975
QY 201 HisArgLysIleSerValGlnTyrGlyAspAlaPheHisProArgProPheCyMetArg 220
Db 976 CACCATAGATCTCGTGTGACAGTACGCGCATGTCTTTCACCCCGGCTTCTGTGATGTGTGT 1035
```

QY 221 ArgValIysGluMetAspAlaIleuValSerProPheAlaArgArgProSerVal 240  
DB 1036 CGCAGTAAAGTGGCAGGTGCTTATGCGATGCTACCTTTGCCGAGCCCGCAGTGC 1095  
QY 241 AsnSerAspLeuIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 260  
DB 1096 AACAGTGAACCTCAAGTCAAGGCGAAGGCTCTACGAGGAGGCGGAGGCGGAGGAGG 1155  
QY 261 AspIysCysThrIleuVal 266  
DB 1156 GACAAAGTGTAGCATCCAG 1173  
RESULT 9  
AF134409 3469 bp mRNA linear ROD 15-JUN-1999  
LOCUS Rattus norvegicus Rhes protein mRNA, complete cds.  
DEFINITION AF134409  
ACCESSION AF134409.1 GI:5059121  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 3469)  
AUTHORS Falk, J.D., Vargiu, P., Foye, P.E., Usui, H., Perez, J., Danielson, P.E.,  
Lerner, D.L., Bernal, J., and Sutcliffe, J.G.  
TITLE Rhes: A Striatum-Specific Ras Homolog Related to Dextral  
J. Neurosci. Res. (1999) In press  
2 (bases 1 to 3469)  
AUTHORS Falk, J.D., Vargiu, P., Foye, P.E., Usui, H., Perez, J., Danielson, P.E.,  
Lerner, D.L., Bernal, J., and Sutcliffe, J.G.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAR-1999) Molecular Biology, The Scripps Research  
Institute, 10550 N. Torrey Pines Road #MB10, La Jolla, CA 92037,  
USA  
FEATURES  
source Location/Qualifiers  
1..3469  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague Dawley CD"  
/db\_xref="taxon:10116"  
/cisue\_type="Brain striatal region"  
852..1652  
/note="similar to Dextral; Ras-like GTP binding protein;  
thyroid hormone dependent"  
CDS  
thyroid hormone dependent  
/codon\_start=1  
/product="Rhes protein"  
/protein\_id="AAD38928.1"  
/db\_xref="GI:5059122"  
/translation="MMKTLSSGNCITLNPAAKNSYRMVVLGASGVKSSIVSFLNGRF  
EDQVPTIEDPHRKYNILHGMVDLIDTSGNHPFPARRSLITLGVFELVSLDS  
RESFDEVRKLOKQILIEVSKLCKMYKKEAEELPMVLCGNKNDSELCROVPANAEALLV  
SGDENCAFEVAKKTNVNEFVLFPSAKLPHMSPLHKLISVQIGDAPHRPFC  
MERTVAGAYGVSPARRPSVNSDLKIKAVLLEGGARERDKCSIQ"  
misc\_feature  
924..947  
/note="Region: GTP-GDP binding site"  
ORIGIN  
PolyA\_signal  
3447..3452  
Alignment Scores:  
Pred. No.: 5..9e-119 Length: 3469  
Score: 1304.00 Matches: 252  
Percent Similarity: 96.24% Conservative: 4  
Best Local Similarity: 94.74% Mismatches: 10  
Query Match: 94.70% Indels: 0  
Gaps: 0  
US-09-778-963A-2 (1-266) x AF134409 (1-3469)  
QY 1 MetMetIysThrIleuSerSerGlyIysCysThrIleuSerValProAlaIysAsnSerVal 20  
DB 852 ATGATGAAAGACCTTGTCCAGCGGAACTGACACTCAACGTCGCCCTAAGAACTCTTAC 911

QY 21 ArgMetValIleuGluAlaSerArgValGlyIysSerSerIleuValSerArgPheIleu 40  
DB 912 CGCATGAGTGTGCTGGTGGCTCCGGGTGGGAAAGACTCATTTGTCTCCGTTCTCT 971  
QY 41 AsnGlyArgPheGluAspGlnThrThrProThrIleuAspPheIleuValIleuVal 60  
DB 972 AATGCGCCCTTTGAGATCAGTACACACCCATCTCGAGGACTTTCATCGCAAGGTGAC 1031  
QY 61 AsnIleuArgGlyAspMetThrGluIleuAspIleuAspThrSerGlyAsnIleuProPhe 80  
DB 1032 AACATCCATGGGAGATGATGAGTGAATCTCGACACCTCTGTAACATCCGTTTC 1091  
QY 81 ProAlaMetArgArgIleuSerIleuThrGlyAspValPheIleuValPheSerIleu 100  
DB 1092 CCGTGCATGCGCCCGGCTGTATCTCATCAGAGAGCGTCTTCATCTGTTTAAAGCTG 1151  
QY 101 AspAsnArgGluSerPheAspGluValIysArgIleuGlnIleuGluValIys 120  
DB 1152 GATAGCCGGGAGTCTTTCAGAGAGGTCAAGGCGCTCCAGAAACAGATCTGAGGTCAAG 1211  
QY 121 SerCysIleuValAsnIleuThrIleuGluIleuValIleuValIleuValIleuVal 140  
DB 1212 TCTGCTGAAGATTAAGACCAAGAGGCGGAGAGCTGCTATGATGATCTGTGGGAAAC 1271  
QY 141 LysAsnAspPheIleuGluIleuValIleuValIleuValIleuValIleuValIleuVal 160  
DB 1272 AAGATGACCAACAGTGAAGTGTGCGGCGGCTCCGTCATGAGGCGGAACTGTGGTG 1331  
QY 161 SerGlyAspGluAsnSerValThrPheGluValSerAlaIleuValAsnValAsp 180  
DB 1332 TCCGGGAGTGAAGAACTGCGCTTCAAGGTGTCAGAGGCAAGAAACAGAACTGAAC 1391  
QY 181 GluMetPheThrValIleuPheSerMetAlaIleuValIleuValIleuValIleuVal 200  
DB 1392 GAGATGTTCTAGTGTCTTTCAGTGGCAGAGCTGCGCCACAGATAGAGCTGTCACATG 1451  
QY 201 HisArgIleuValSerValGlnThrIleuValAspIleuValIleuValIleuValIleuVal 220  
DB 1452 CACCAAGATCTCGTGCAGTACAGTACGCGCTTCCACCCCGGCTTCTGATGTCGT 1511  
QY 221 ArgValIysGluMetAspAlaIleuValIleuValIleuValIleuValIleuValIleuVal 240  
DB 1512 CGCACCAAGTGGCAGGTGCTATGCGTGTCTGCGCTTCCGCGCGGCGCCAGTGC 1571  
QY 241 AsnSerAspLeuIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 260  
DB 1572 AACAGTGAACCTCAAGTCAAGGCGAAGTCTTACGAGGAGGCGGAGGCGGAGGAGG 1631  
QY 261 AspIysCysThrIleuVal 266  
DB 1632 GACAAAGTGTAGCATCCAG 1649  
RESULT 10  
HSM803172 2699 bp mRNA linear PRI 10-JUL-2002  
LOCUS Homo sapiens mRNA; cDNA DKFZp761D0217 (from clone DKFZp761D0217).  
DEFINITION AL831936  
ACCESSION AL831936  
VERSION AL831936.1 GI:21732458  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2699)  
AUTHORS Anseorge, W., Winkler, U., Mewes, H.W., Weil, B. and Wiemann, S.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuberg, GERMANY  
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;  
sequenced by EMBL (European Molecular Biology Laboratories,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.



This clone (DKFZp761D0217) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcententrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cDNA/>.

## FEATURES

source

1..2699

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="RZPD:DKFZp761D0217"

/db\_xref="taxon:9606"

/map="22q13.1"

/clone="DKFZp761D0217"

/issue\_type="amygdata"

/clone\_id="761 (synonym: hamy2). Vector pSPORT1; host

DH10B; sites NotI + SalI"

/dev\_stage="adult"

2660..2665

2677

polya\_signal

polya\_site

## ORIGIN

polya\_signal

2677

## Alignment Scores:

Pred. No.:	6,81e-109	Length:	2699
Score:	1201.00	Matches:	235
Percent Similarity:	98.74%	Conservative:	0
Best Local Similarity:	98.74%	Mismatches:	3
Query Match:	87.22%	Indels:	1
		Gaps:	0

US-09-778-963A-2 (1-266) x HSM803172 (1-2699)

Qy 29 ArgValGlyLysSerSerIleValSerArgPheLeuAengIYAArgPheGluAAspGlnTyr 48  
Db 2 CGAGTGGGCAAGACCTCATGCTGCTGCTTCCTCAATGAGCCGCTTGAGGACCAAGTAC 61  
Qy 49 ThrProThrIleGluAAspPheHISArgLysValTyrAsnIleArgGlyAAspMetTyrGln 68  
Db 62 ACACCCACCATCGAGACTTCCACCGTAAAGTATTCACATCCGGGCGGACATGATCAG 121  
Qy 69 LeuAspIleLeuAAspThrSerGlyAAsnHisProPheProAlaMetArgArgLeuSerIle 88  
Db 122 CTCGACATCCTGGATACCTCTGGCAACCACTCCCTCCGATGCGGATGCGATGCTCATC 181  
Qy 89 LeuThrGlyAAspValPheIleLeuValPheSerLeuAAspAAsnArgGlySerPheAspGlu 108  
Db 182 CTCACAGGGGATGCTTCATCTGCTGTTCAAGCCGATTAACCGGAGTCTCTTCATAG 241  
Qy 109 ValIysArgLeuGlnLysGlnIleLeuGlnValLysSerCysLeuLysAAsnLysThrIys 128  
Db 242 GTCAAGCGCTTTCAGAACCAATCTGAGAGTCAATCTCTGCTTAAGAACCAAGCCAG 301  
Qy 129 GluAlaIaGluLeuPheMetValIleCysGlyAAsnLysAAsnAspHisGlyGluLeuCys 148  
Db 302 GAGGGGGGAGAGCTGCCATGCTATCTGTGGCAACAAAGACGACGAGGAGCTGTGC 361  
Qy 149 ArgGlnValProThrThrGluAlaGluLeuValLysSerGlyAAspGluAAsnSerAlaTyr 168  
Db 362 CGCCAGGCGCCACACCAAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
Qy 169 PheGluValSerAlaValLysAAsnThrAAsnValAAspGluMetPheTyrValLeuPheSer 188  
Db 422 TTCAGAGTGTGGCCAAAGAACCAACGAGGAGAGATGTTTACGATGCTCTTCAGC 481  
Qy 189 MetAlaLysLeuProHisGluMetSerProAlaLeuHisArgLysIleSerValGlnTyr 208  
Db 482 ATGGCAAGCTGCCACAGATGAGCCCGCCCTGCATCGAATATCTCGTGCAGTAC 541  
Qy 209 GlyAAspAlaPheHisArgProPheCysMetAAsnGlnValLysGluMetAspAlaTyr 228  
Db 542 GGTGACGCTTCCACCCAGGCGCTTGTGCATGCCCGGCTCAAGAGATGAGACCCCAT 601  
Qy 229 GlyMetValSerProPheAlaArgProSerValAAsnSerAAspLeuLysTyrIleLys 248

Db 602 GCATGCTCGCCCTCGCCCGCCCGCCGCTCAACAGTACCTCAATTACATCAAG 661  
Qy 249 AlAysValLeuAAspGluGlnAlaArgGluAAspLysCysThrIleGln 266  
Db 662 GCCAAGGCTCTCGGAGAGGCAAGCC-CGTGAGAGGAGCAAGTGCACATTCAG 714  
RESULT 11  
AC145933  
LOCUS  
DEFINITION  
GALLUS gallus chromosome UNK clone CH261-41E11, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 8 unordered pieces.  
AC145933  
VERSION  
AC145933.1 GI:33386883  
KEYWORDS  
HTG: HTGS PHASE1.  
SOURCE  
Gallus gallus (chicken)  
ORGANISM  
Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 210264)  
Wilson, R.K.  
The sequence of Gallus gallus clone  
Unpublished  
2 (bases 1 to 210264)  
Wilson, R.K.  
Direct Submission  
Submitted (01-AUG-2003) Genetix, Genome Sequencing Center, 4444  
Forest Park Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 210264)  
Wilson, R.K.  
Direct Submission  
Submitted (19-NOV-2003) Genetix, Genome Sequencing Center, 4444  
Forest Park Parkway, St. Louis, MO 63108, USA  
COMMENT  
Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Project Information  
Center project name: J\_A041B11  
----- Summary Statistics -----  
Sequencing vector: M13, 0%

Chemistry: Dye-Primer ET; 0% of reads  
Chemistry: Dye-Terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 207188 bases at least Q40  
Consensus quality: 207951 bases at least Q30  
Consensus quality: 208354 bases at least Q20  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1093: contig of 1093 bp in length  
\* 1094 1193: gap of unknown length  
\* 1194 2207: contig of 1014 bp in length  
\* 2208 2307: gap of unknown length  
\* 2308 24970: contig of 22663 bp in length  
\* 24971 25070: gap of unknown length  
\* 25071 48476: contig of 23406 bp in length  
\* 48477 48577: gap of unknown length  
\* 48578 76319: contig of 27643 bp in length  
\* 76320 76320: gap of unknown length  
\* 141133: contig of 64814 bp in length  
\* 141234 208929: gap of 67696 bp in length  
\* 208930 209029: gap of unknown length

```

FEATURES          * 209030 210264: contig of 1235 bp in length.
source             1..210264
                  /organism="Gallus gallus"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9031"
                  /chromosome="UNK"
                  /clone="CH261-41E11"
misc_feature       1..1093
                  /note="assembly_name:Contig10"
misc_feature       1194..2207
                  /note="assembly_name:Contig11"
misc_feature       2308..24970
                  /note="assembly_name:Contig12"
misc_feature       25071..48476
                  /note="assembly_name:Contig13"
misc_feature       48577..76219
                  /note="assembly_name:Contig14"
misc_feature       76320..141133
                  /note="assembly_name:Contig15"
misc_feature       141234..208929
                  /note="assembly_name:Contig16"
misc_feature       209030..210264
                  /note="assembly_name:Contig9"

ORIGIN
Alignment Scores:
Pred. No.:         2,83e-82          Length:         210264
Score:             955.50            Matches:         239
Percent Similarity: 32.19%           Conservative:    14
Best Local Similarity: 30.41%        Mismatches:      13
Query Match:       69.39%            Indels:          520
DB:                2                 Gaps:            1

US-09-778-963a-2 (1-266) x AC145933 (1-210264)
QY      1 MetMetLySfThLeuSerSerGlyAenCySfThLeuSerValProAlaLysAsnSerTyr 20
Db      36607 ATGATGAGAGACCATGCTCTAGTGGAACCTGCACCTTAATGTCCAGCCAAAGACTCGTAC 36666
QY      21 ArgMetValValLengIValaSerArgValGlyLysSerSerIleValSerArgPheLeu 40
Db      36667 CCGATGAGTGTGTGGAGCCCTCCAGGGGTGGGAAAAGTTCCATATGCTCACGCTTCTC 36726
QY      41 AenGlyArgPheGlyuAspGlyntYrThrProThrIleGlyuAspPheHisArgLysValTyr 60
Db      36727 AATGGCCGCTTTGAGGACCAAGTACACTCCACCATTTGAGGATTTTCCACCGCAAGCTTAC 36786
QY      61 AenIleArgGlyAAspMetYrGlnLeuAepIleLeuAepThrSerGlyAAsnHisProPhe 80
Db      36787 AACATCCCGGGAGATATGTATGATGAGCTGAGCATCTGAGACCTCAGGGAATCACCTTTTC 36846
QY      81 ProAlaMetArgArgLeuSerIleLeuThr----- 90
Db      36847 CCTGCTATAGAGAGGCTTTCCATCTCTGACAGGTGAGAAAGTAAAGAAAGTTGCACAAACAG 36906
QY      90 ----- 90
Db      36907 GGAACAAGAGATAAAGATTATATTAATTAACAATGTAGCTTGAAAACTGTATCTGTGAAGT 36966
QY      90 ----- 90
Db      36967 GAGGTGTTCTTGGGATATAGATGCAATTGGGAGGACCAAAAATGGGCTGTGGGATTTTCAT 37026
QY      90 ----- 90
Db      37027 TGCAAGATTCACTTCTCCAGAGAGATGTTTCTGTGGCTGTATGCTTCTATCTGTACAG 37086
QY      90 ----- 90
Db      37087 AATGACTAGCCACAGAGAGAACTCACTGGGTGAAGGTGAAATTAAGATTTCCCCC 37146
QY      90 ----- 90

```

```

Db      37147 TAGTTATATTCGCCGCTGTGCATTTCCCTCTTTCGAACATATTTTTTCTCGCCGTCA 37206
QY      90 ----- 90
Db      37207 CAATGAATGTTGATATTAATCTGTGGCAATTAAGAGCAGCAAGGAGACAAAACAGGC 37266
QY      90 ----- 90
Db      37267 ATGTCAATTATAGACACAGAGCATGAATAATGAGTGAATATGCACAGGAGGAGTTAA 37326
QY      90 ----- 90
Db      37327 AACTGAGAAAGCTCAAGATTTTTTTCTTTCTTAAGTCCAGCCGTGATCCCTTGA 37386
QY      90 ----- 90
Db      37387 TTTAATATTCAGTAAAAAAGATCACTGTACTTTTACCTTCTCTTAAGTCAAAAT 37446
QY      90 ----- 90
Db      37447 AGTCAATGAATTAAGAGGAAATAGCGGACTTTTGTACATCCCATTTGTTAGT 37506
QY      90 ----- 90
Db      37507 GGTTCCTGTGTGAGCAAAAGATATGTCTTAACGTAAAGTGACAAATGAGATATATA 37566
QY      90 ----- 90
Db      37567 TTCAACACAGCTACAAACAAACAACTTTGCGTGTGTCTTCCCTCTAGCTAGC 37626
QY      90 ----- 90
Db      37627 CAGTGAACATGACATTGTATCAGATGGGCTACCATCTGCATTATGAGAGTCTTAG 37686
QY      90 ----- 90
Db      37687 GCCAGTTTCTGGACATGATGTAGCTTTTGTCCATGAAGTCTTTCACAGTTTCAACA 37746
QY      90 ----- 90
Db      37747 TGACTGTATTCACAGGCAATGATTTCTTTGCAAAAGAAATGTATGTCAATAAATGTC 37806
QY      90 ----- 90
Db      37807 AGAAGAAATCCAGGGAATCTGTGTCAATTATGTATTTCTTTTCCATGTCGTGTT 37866
QY      90 ----- 90
Db      37867 CTCTAAGAAATGAATATGACAGACAGACAGTGAACATTTGGAAAGCTCTCATCTCTCAC 37926
QY      90 ----- 90
Db      37927 TTAGGCAATTTGGAGCTGACTAGCTGGAATCTCTCATTTCAAGTGTTTCTTTGTATTAAC 37986
QY      90 ----- 90
Db      37987 ATAAATCAGAGGCCAAGGCTAAGTGAATTTAGTTTCCACTTCTCTGTGTAAGACTA 38046
QY      90 ----- 90
Db      38047 GCTGTGTTTACACAGCAATACAGACAAATGAAGAAAAAGAAAGCAATGCACAGAGC 38106
QY      90 ----- 90
Db      38107 TGGATAGAACTGTGTTTACAAAGACAGTACAAATGGGCAATAGCTGCAGTGC 38166
QY      90 ----- 90
Db      38167 AACCTAGGGGTACATCTCAGATATGGCTGAATATACATCAAGCCATAAGCAATTTGTAC 38226
QY      90 ----- 90

```

```

Db 38227 TACAGAAATGAGTGGGAGTTGGATTAGAGAACTGAAACTAAAGTAAACAGATGAAAG 38286
Qy 90 -----
Db 38287 AGATTCAAAAGACACACAGAAAAAATTAAAAAATGAACAACATGTGGGAGG 38346
Qy 90 -----
Db 38347 AGACAGCAATATAATGATATGTTTGTGAAAATTGGCTTCTCTGGCACAGGTTGCTGAA 38406
Qy 91 -----
Db 38407 GCATTTCAGTTCCTTTTTCATTCAGATTCAGCGGATGTTTCCTCCGATTCAGCTTGA 38466
Qy 101 PAAATGAGTTCCTTTTTCATTCAGATTCAGCGGATGTTTCCTCCGATTCAGCTTGA 38466
Db 38467 TAAACAGAGATCTCTTTGATGAGTCAAGAGACTCCAGAGAGAGATCTCTCGAGGTTAAATC 38526
Qy 121 TCTGCTGAGAGAAACAGACCAAGAGATCAGCTGACCTTCCATGATCTGTGGCAACA 38586
Db 38527 CTGGCTGAGAGAAACAGACCAAGAGATCAGCTGACCTTCCATGATCTGTGGCAACA 38586
Qy 141 SASATGAGTTCCTTTTTCATTCAGATTCAGCGGATGTTTCCTCCGATTCAGCTTGA 38646
Db 38587 AAATGACCAACAGTAAATCTACCGCAAGTCCGCTCAGATGAAAGGTTGAAACCTTGTGC 38646
Qy 161 TGLYASPLUSNSERIALTYRPHGLVALSERIALYSLYASNTTAAVALASPGI 181
Db 38647 CAGTGATGAAAGTCCGCTTCTGAGGTTGAGCTTAAAGAAACACCAACGATGATGA 38706
Qy 181 UMETPHETVALLEUPHESERMETALALYSLUPROHISGLUMETSERPROALALUHI 201
Db 38707 GATTTTCATGTTCTCTTCCAGATGCGCAAGTACTCATGATGAGACCTTCCCTTCCA 38766
Qy 201 SARGULYSERVALGINTYRGYASPALPHENISPROARGPROPECYMETARGAR 221
Db 38767 CAGGAAATCTCCATCCAGTACGGGACACCTTCCCAAGAAATCTTCCGATGTCGACG 38826
Qy 221 GVALYGLUMETAPALATYRGLYMETVALSERPROPHETALARGAARGPROSERVALAS 241
Db 38827 AGTAAAGACATGATGATGCTTATGATATCTCTCTCTGCTGCGCCAGAGGTCAA 38886
Qy 241 NSEASPLEULYSTYRILEYALALYVALLEUARGLUGLUGLVALARGGLUARGAS 261
Db 38887 CAGGACCTGAGATGATCAATCAAAAGTTCTCAGGAGAGTCACTCAAGGAGAGGGA 38946
Qy 261 PLYSCYSTRILEGIN 266
Db 38947 GAAATGACATCCAG 38962

```

```

RESULT 12
LOCUS BC056272 1390 bp mRNA linear VRT 08-OCT-2003
DEFINITION Danio rerio cDNA clone MGC:65909 IMAGE:6788683, complete cde.
ACCESSION BC056272
VERSION BC056272.1 GI:33989992
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 1390)
Strausberg, R.L., Feltingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Uesdin, T.B., Toshimiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Mortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Hellon, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bonfield, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
12477932
2 (bases 1 to 1390)
Strausberg, R.
Direct Submission
Submitted (11-AUG-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bonfield, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Madraro, Q.L., Masello, C., Maskeri, B., Mastriani, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stancirpop, S., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 121 Row: a Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. 1390
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:65909 IMAGE:6788683"
/tissue_type="Embryo, zebrafish 24 hpf"
/clone_lib="NCI CGAP_Zemb2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
35..832
/codon_start=1
/product="Unknown (protein for MGC:65909)"
/protein_id="AAH56272.1"
/db_xref="GI:33989993"
/translation="MTKMSPSSENEPIIPAKNCYRMTLSTKYGKTAIYRFLNGR
RESDYPTIEDFHRKLYSIKGDVYQDLIDTSGNHPFAMRLSLITDVFILVSLDN
RESFHEVORLKQOIVETKSGCKNKTENVDVPIVTCNGKGDREFRYEVDREISOLA
KDEGCAYFEISAKNTVDQFQRLFTLAKIPNMSPLDRKVSVOYCDILHKSLKN
KXVDGAYGIVAPFARSPVSHDMLYIKERAKIGGQGTOKERCITIS"
86..604
/note="RAS; Region: Ras subfamily of Ras small GTPases"
/db_xref="CCD:smart00173"
ORIGIN
misc_feature
Alignment Scores:
Pred. No.: 1.17e-83 Length: 1390

```

Score: 942.50 Matches: 177  
 Percent Similarity: 83.02% Conservative: 43  
 Best Local Similarity: 66.79% Mismatches: 44  
 Query Match: 68.45% Indels: 1  
 DB: 5 Gaps: 1

US-09-778-963A-2 (1-266) x BC056272 (1-1390)

```

QY 1 MetMetlyThrLeuSerSerGlyValnCySerThrLeuSerValProAlaIleAsnSerTyr 20
DB 35 ATGATCAAGAAATATGATCCGTCGGAACAAAGATTGACATCCCGCGAAGAAATCGCTAC 94
QY 21 ArgMetValIleLeuGlyAlaSerArgValGlyLeuSerSerIleValSerArgPheLeu 40
DB 95 CGGATGGATGCTCTGGGTCACCAAGTCGGGAGAGACGGCATCTGTGCGCTTTCTG 154
QY 41 AsnGlyArgPheGluAspGlnTyrThrProThrIleGluAspPheHisArgIleValTyr 60
DB 155 AACGACCGCTTTGAGAGACAGTACAGCGCATTTAGGACTTTGATAGAACTCTAC 214
QY 61 AsnIleArgGlyAspMetTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 80
DB 215 AGCATTAAGAGAGATGATGATCAGCTGATATTTGGATCTTCACTTCAGGGAAACCATCATTC 274
QY 81 ProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeu 100
DB 275 CCAGCTATGCGGAGACTGTCATCTGACCGGATGATGATTTATCTGCTGCTGCTGCTGCTG 334
QY 101 AspAsnArgGluSerPheAspGluValIleAspArgLeuGlnIleLeuGluValIle 120
DB 335 GACACACCGGAGATCTCTCCACAGAGGTGACGCTGACAGACAGACAGATCTACAGACACAA 394
QY 121 SerGlyLeuValAsnIleThrIleValIleAspGlnLeuPheMetValIleCysGlyAsn 140
DB 395 TCTCTCCCTTAAGAAACAAAGACCAAGAGAACTGACAGCTGCTGCTGCTGCTGCTGCTG 454
QY 141 LysAsnAspPheIleGlyLeuValIleProThrThrGlnValIleGluLeuVal 160
DB 455 AAGGCGGACCGC--GAGTTTACCGAGAGTTCACCGGACAGACAGATGACAGACGCTGATC 511
QY 161 SerGlyAspGluAsnSerIleThrPheGluValIleSerIleValIleAsnThrAsnValAsp 180
DB 512 GCCGAGACAGACAAATGCGCTACTTCGAGATCTCCGCTAAAGAAACGAAACGCTGATC 571
QY 181 GluMetPheTyrValLeuPheSerMetAlaIleValLeuProHisGluMetSerProAlaLeu 200
DB 572 CAGATGTTCCAGAGACTCTTTACGCTCGCCAAACCTGCCAATGAGATGAGCCCTGATCTC 631
QY 201 HisArgIleValSerValGlnTyrGlyAspAlaPheHisProArgProPheCysMetArg 220
DB 632 CATCCGAAAGGTCTCCGTGACATCTGATCATCTCCATMAAAAGTCGCTGAAGAAACAG 691
QY 221 ArgValIleValMetAspAlaTyrGlyMetValIleSerProPheAlaArgProSerVal 240
DB 692 AAGGTGAAGACGAGCGGACGCTGATTCGCGCTTTGCGCGGCGCGCGCGCGCGCG 751
QY 241 AsnSerAspLeuValTyrIleValValIleValIleValIleValIleValIleValIle 260
DB 752 CACATGTGATTAATGATCAAGAAAGAGATAGAGAGTGAAGAGCTAAGATTA 811
QY 261 AspLysCysThrIle 265
DB 812 GAAAGGTGATTAATC 826

```

RESULT 13  
 LOCUS AX572850 11221 bp DNA linear PAT 29-NOV-2002  
 DEFINITION Sequence 3 from Patent WO02062849.  
 ACCESSION AX572850  
 VERSION AX572850.1 GI:26004938  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Neelam, B., Ketchum, K.A., di Francesco, V., and Beasley, E.M.  
 Isolated human ras-like proteins, nucleic acid molecules encoding  
 them, and uses thereof  
 Patent: WO 02062849-A 3 15-AUG-2002;  
 JOURNAL PE Corporation (NY) (US)

FEATURES  
 source location/Qualifiers

1..11221  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

#### ORIGIN

##### Alignment Scores:

Pred. No.: 1.98e-79 Length: 11221  
 Score: 911.00 Matches: 176  
 Percent Similarity: 98.33% Conservative: 1  
 Best Local Similarity: 97.78% Mismatches: 3  
 Query Match: 66.16% Indels: 0  
 DB: 6 Gaps: 0

US-09-778-963A-2 (1-266) x AX572850 (1-11221)

```

QY 87 SerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgIleSerPhe 106
DB 7681 TCTCTCCCTCGAGGAGATGCTCTCATCTGATGTTCACTCGATACCTCGATACCGGAGTCTTC 7740
QY 107 AspGluValIleArgLeuGlnIleValIleLeuGluValIleValIleValIleValIle 126
DB 7741 GATGAGGTCAAGCGCTTCAGAGACAGATCTCGAGGATCAAGTCTGCTGCTGCTGCTGCTG 7800
QY 127 ThrIleValIleAlaIleGluLeuProMetValIleCysGlyAsnIleAsnAspHisGlyVal 146
DB 7801 ACCAAGAGGCGGCGGAGACTCCCATGCTCATCTGTGGCAACAAAGACACGACCGCGAG 7860
QY 147 LeuCysArgGlnValProThrThrGlnIleGluLeuValIleSerGlyAspGluAsnSer 166
DB 7861 CTGTGCGCCGAGGTGCCACACACGAGGCCGAGCTGCTGTGCTGCTGCTGCTGCTGCTG 7920
QY 167 AlaTyrPheGluValIleSerIleValIleValIleValIleValIleValIleValIle 186
DB 7921 GCTTACTTCGAGGTGTGCGCCAAAGAAACACCAACGTCGACAGATGTTCTTACGTCCTC 7980
QY 187 PheSerMetAlaIleValLeuProHisGluMetSerProAlaIleHisArgIleIleSerVal 206
DB 7981 TTCAGCATGCGCCAAAGCTTCCACACAGAGATGAGCCCGCTGTCATCGCAAGATCTCCGTG 8040
QY 207 GlnTyrGlyAspAlaPheHisProArgProPheCysMetArgArgValIleGluMetAsp 226
DB 8041 CAGTACGATGACGCTTTCACACCCAGGCGCTTCTGCACTGCGCGGTCAAGAGATGAGAC 8100
QY 227 AlaTyrGlyMetValIleSerProPheAlaArgProSerValIleAsnSerAspLeuIleTyr 246
DB 8101 GCTTATGACATGATGTCTGCGCTTTCGCGCGCGCCCAACGTCACAGATGACCTCAAGTAC 8160
QY 247 IleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 266
DB 8161 ATCAAGGCGCAAGGTCTTTCGGAAGGCCAGGCGCTGAGAGGAGCAAGTGCACCATCCAG 8220

```

RESULT 14  
 LOCUS HS569D19 114771 bp DNA linear PRI 05-JUN-2003  
 DEFINITION Human DNA sequence from clone RPA-569D19 on chromosome 22q13.1,  
 complete sequence.  
 ACCESSION AL022334  
 VERSION AL022334.1 GI:3947839  
 KEYWORDS HTG  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 114771)  
AUTHORS Williams, S.  
TITLE Direct Submission  
JOURNAL Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

## COMMENT

On Dec 2, 1998 this sequence version replaced gi:3445463.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep -----  
Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquerry@sanger.ac.uk

-----  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.  
This sequence was generated from part of bacterial clone contigs of  
human chromosome 22, constructed by the Sanger Centre Chromosome 22  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr22  
RP4-569D19 is from the library RPCI-4 constructed by the group of  
Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCYPAC2  
IMPORTANT: This sequence is not the entire insert of clone  
RP4-569D19 it may be shorter because we sequence overlapping  
sections only once, except for a short overlap.  
The true left end of clone RP4-569D19 is at 1 in this sequence. The  
true right end of clone RP5-824119 is at 31286 in this sequence.  
Location/Qualifiers

FEATURES  
source  
1..114771  
/organism="Homo sapiens"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/map="q13.1"  
/clone="RP4-569D19"  
/clone\_1id="RPCI-4"  
142..175  
/note="MIR3 repeat: matches 162. .195 of consensus"  
277..290  
/note="2.3 copies 6 mer GGAGAA 28% conserved"  
418..427  
/note="3.3 copies 3 mer CTC 20% conserved"  
514..524  
/note="2.2 copies 5 mer CAGCA 22% conserved"  
756..765  
/note="2.5 copies 4 mer CCCA 20% conserved"  
complement(805..1011)  
/note="MIR repeat: matches 50. .258 of consensus"  
1012..1253  
/note="AluSc repeat: matches 61. .304 of consensus"  
complement(1466..2088)  
/note="L2 repeat: matches 2487. .3183 of consensus"  
2302..2611  
/note="AluSg repeat: matches 1. .309 of consensus"  
2616..2630  
repeat\_region

repeat\_region /note="3.8 copies 4 mer AAAC 30% conserved"  
2691..2706  
/note="2.7 copies 6 mer GTGGGA 25% conserved"  
2748..3041  
/note="L2 repeat: matches 3033. .3313 of consensus"  
3546..3871  
/note="AluSx repeat: matches 1. .312 of consensus"  
3886..4117  
/note="MIR repeat: matches 13. .237 of consensus"  
complement(4206..4486)  
/note="Aluub repeat: matches 1. .282 of consensus"  
complement(4747..4858)  
/note="MIR repeat: matches 47. .163 of consensus"  
4882..4959  
/note="L2 repeat: matches 3187. .3266 of consensus"  
5079..5101  
/note="1.9 copies 12 mer AATTCATTGATG 46% conserved"  
5080..5103  
/note="6.0 copies 4 mer ATTC 30% conserved"  
5108..5568  
/note="match: GSS: Em:AA520278"  
5219..5229  
/note="2.8 copies 4 mer TCTT 22% conserved"  
5324..5334  
/note="2.2 copies 5 mer CTCGC 22% conserved"  
5493..5545  
/note="2.1 copies 25 mer GAAGGAGAGCTTTATTCAGCA 88% conserved"  
5693..5990  
/note="AluY repeat: matches 1. .300 of consensus"  
5999..6140  
/note="AluSg/x repeat: matches 1. .136 of consensus"  
6147..6202  
/note="MIR repeat: matches 151. .205 of consensus"  
complement(6238..6292)  
/note="MIR3 repeat: matches 147. .203 of consensus"  
6330..6339  
/note="10.0 copies 1 mer T 20% conserved"  
6342..6352  
/note="2.2 copies 5 mer GTGGG 22% conserved"  
6453..7182  
/note="L1M4S repeat: matches 4129. .4859 of consensus"  
complement(7294..7605)  
/note="AluSx repeat: matches 1. .312 of consensus"  
7630..7640  
/note="2.8 copies 4 mer ATTA 22% conserved"  
7786..8087  
/note="AluSg repeat: matches 6. .307 of consensus"  
8748..8909  
/note="AluY4S2 repeat: matches 149. .310 of consensus"  
8910..8965  
/note="28.0 copies 2 mer AT 112% conserved"  
9217..9231  
/note="3.0 copies 5 mer GCTTT 21% conserved"  
complement(9458..9902)  
/note="match: GSS: Em:AA222044"  
complement(9546..9904)  
/note="match: GSS: Em:B57571"  
10091..10506  
/note="L1M4S repeat: matches 4845. .5273 of consensus"  
10508..10592  
/note="AluY/FLM repeat: matches 1. .88 of consensus"  
10593..11608  
/note="L1M4S repeat: matches 5265. .6298 of consensus"  
11640..11653  
/note="14.0 copies 1 mer A 28% conserved"  
complement(11931..12080)  
/note="L1M41 repeat: matches 6010. .6161 of consensus"  
12092..12108  
/note="2.4 copies 7 mer TGTTTC 25% conserved"  
12150..12164  
/note="7.5 copies 2 mer AG 23% conserved"  
complement(12342..12784)  
repeat\_region



Search completed: June 20, 2004, 22:13:28  
Job time : 4312 secs

\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 49224: contig of 49224 bp in length  
\* 49225 49324: gap of unknown length  
\* 49325 49327: contig of 16013 bp in length  
\* 49327 49338 65437: gap of unknown length  
\* 49338 65438 73997: contig of 8560 bp in length  
\* 65438 73998 74097: gap of unknown length  
\* 73998 74098 129526: contig of 55429 bp in length  
\* 74098 129527 129626: gap of unknown length  
\* 129527 129627 178071: contig of 48445 bp in length.  
Location/Qualifiers  
1. 178071  
/organism="Lemur catla"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9447"  
/clone="LB2-219P16"

## FEATURES

## ORIGIN

## Alignment Scores:

Pred. No.:	1.63e-77	Length:	178071
Score:	906.50	Matches:	177
Percent Similarity:	93.75%	Conservative:	3
Best Local Similarity:	92.19%	Mismatches:	11
Query Match:	65.83%	Indels:	1
DB:	2	Gaps:	1

US-09-778-963A-2 (1-266) x AC145536 (1-178071)

QY	76	GLYASNHISPRPHEPROALMETARGRGLUSERILEU---THRGLYASPYALPHE	94
DB	59153	GGGTCGAGCGAGCCCGCGCTTCCTCCCGCAGCGACGCTCTTC	59212
QY	95	ILEUVALPHESERLAUSPASNARGIUSERPHEASPGIUVALLYARGLEUGLNYLS	114
DB	59213	ATCCTGGTGTTCACGCTCGATACCGGAGTCTTCGACGAGTCAAGCGCTGCAAG	59272
QY	115	GLNILEUGIUVALLYSERCYSLAUSANLYSTHRLYSGIUALAIGLULEUPRO	134
DB	59273	CAGATCTCGAGGTCAAGTCTGCTTAAGAAACAAGCAAGAGGCGAGCTGCC	59332
QY	135	METVALILECYSGIYASNLVYASNAPHSISGLYGLULEUCYARGINVALPROTHR	154
DB	59333	ATGCTCATCTGCGGCAACMAAGACGCGAGCTCTGCCGAGTCCCAACACC	59392
QY	155	GLUALAGIULEUVALSERGIYASPGIUAHNSERIALATYRPHEGLUVALSERALAY	174
DB	59393	GAGGCCGAGCTGCTGGTCTCAGGGGACGAACTGCCCTTCTCGAGGTGTCCGC	59452
QY	175	LYSASNTHRAENVALASPGIUMETPHEIYRVALLEUPHESEMETALALYSLEUPRO	194
DB	59453	AAGAACAACCAAGTGAATGATGTTTACGTGCTTCAAGATGCGCAAGCTGCCG	59512
QY	195	GLUETSERPROALAEUHSIARGULYSLESERVALGINIYRGLYASPAALPHEHIS	214
DB	59513	GAGATGAGCCCCGCGCTGCAACGATCTCTGTGACGTACGGGACGCTTCCACCC	59572
QY	215	ARGPROPHCYSMETARGARGVALYSGIUMETASPAALATYRGIMETVALSERPRO	234
DB	59573	AGGCCCTTCTGCAATGCGCGCATCAAGAGGCGCAGCTTACGGCATGTCTTCC	59632
QY	235	ALAHARGPROSERVALASNSEASPLEULYSTYRIELYSAIALYSVALLEUARG	254
DB	59633	GCCCGCGCGCCCAAGTCAACAGTGAACCTCAAGTACATCAAGCCCAAGTCTCC	59692
QY	255	GLYGLINALAARGIUALARGAPLYGCTHRILEGIN	266
DB	59693	GGCCAGGCGCTGAGAGGGAACAATGACCATCCAG	59728

THIS PAGE BLANK (USPTO)





Db	181	GGCCCGGCCCAATCCAGCCCCGAGCCATATGAACCTTTGTCCAGCGGAACCTC	240
Qy	241	ACGCTCAGTGTGCCCGCCAAAATCTCATCCGATGATGTGTGCTTGGGTCTCTCGGGTG	300
Db	241	ACGCTAGTGTGCCCGCCAAAATCTCATCCGATGATGTGTGCTTGGGTCTCTCGGGTG	300
Qy	301	GGCAAGAGCTTCATCTGTCTCTGCTTCTCTATATGCGCGCTTTAAGAACCAATACACCC	360
Db	301	GGCAAGAGCTTCATCTGTCTCTGCTTCTCTATATGCGCGCTTTAAGAACCAATACACCC	360
Qy	361	ACCATCAGAGACTTCCACCGTAAGTATACAACTCCGCGCCACATGTATCCAGCTGCAC	420
Db	361	ACCATCAGAGACTTCCACCGTAAGTATACAACTCCGCGCCACATGTATCCAGCTGCAC	420
Qy	421	ATCTGATACCTCTGGCAACCAACCCCTTCCCGCCATGCGCAGGCTGTTCATCTTCACA	480
Db	421	ATCTGATACCTCTGGCAACCAACCCCTTCCCGCCATGCGCAGGCTGTTCATCTTCACA	480
Qy	481	GGGATATCTTCATCTGTGTGTAGCGCTGGATTAACGGGAGTCTTCATGAGGTCAAG	540
Db	481	GGGATATCTTCATCTGTGTGTAGCGCTGGATTAACGGGAGTCTTCATGAGGTCAAG	540
Qy	541	CGCCTTCAGAAAGCAGATCTTGAAGTCAAGTCTCTGCTTAAGAACCAAGACCAAGAGCG	600
Db	541	CGCCTTCAGAAAGCAGATCTTGAAGTCAAGTCTCTGCTTAAGAACCAAGACCAAGAGCG	600
Qy	601	GGGAGCTGCCATGTGTATCTGTGGCAAGAAAGAACCAACCGCGAGCTGTGCGCCAG	660
Db	601	GGGAGCTGCCATGTGTATCTGTGGCAAGAAAGAACCAACCGCGAGCTGTGCGCCAG	660
Qy	661	GTGCCACACACCGAGCGAGCTGTGTGTGTGGGAGACAGAACTTCGCTTACTTTCAG	720
Db	661	GTGCCACACACCGAGCGAGCTGTGTGTGTGGGAGACAGAACTTCGCTTACTTTCAG	720
Qy	721	GTGTGGCCAGAGAAACCAACGATGACAGATGTTTATACGTGTCTTCAAGATGGCC	780
Db	721	GTGTGGCCAGAGAAACCAACGATGACAGATGTTTATACGTGTCTTCAAGATGGCC	780
Qy	781	AAGCTGCCACAGAGATGAGCCCGCTGTGATCGCAAGATCTCCGTGAGTACCGGTAC	840
Db	781	AAGCTGCCACAGAGATGAGCCCGCTGTGATCGCAAGATCTCCGTGAGTACCGGTAC	840
Qy	841	GCCTTTCACCCAGGCGCTTCTGCATGCGCGCGCTCAAGAGATGACGCTATGGCATG	900
Db	841	GCCTTTCACCCAGGCGCTTCTGCATGCGCGCGCTCAAGAGATGACGCTATGGCATG	900
Qy	901	GTCTGSCCTTGGCCCGCGCCGACGCTCAAGATGACCTCAAGTACATCAAGAGCCAG	960
Db	901	GTCTGSCCTTGGCCCGCGCCGACGCTCAAGATGACCTCAAGTACATCAAGAGCCAG	960
Qy	961	GTCTTCGGAAAGCCAGGCGCGGTGAGAGGACAAATGCACATTCAGTGAAGGAGAT	1020
Db	961	GTCTTCGGAAAGCCAGGCGCGGTGAGAGGACAAATGCACATTCAGTGAAGGAGAT	1020
Qy	1021	GCTGGGCGGGGCTTGGCAGTGGCTTCAAGGAGGTGGCCCAATGCGCACTGTGGCA	1080
Db	1021	GCTGGGCGGGGCTTGGCAGTGGCTTCAAGGAGGTGGCCCAATGCGCACTGTGGCA	1080
Qy	1081	TCTCCCAACGAGGCGCCGGACGAGTCTTGTTCACAGACCTTAGGCAACGACTGAGG	1140
Db	1081	TCTCCCAACGAGGCGCCGGACGAGTCTTGTTCACAGACCTTAGGCAACGACTGAGG	1140
Qy	1141	CCCCCGGGGCTGGCTTCGCAATTCGTCTGCTTTCACAGCTTTCAGATGCGGCTT	1200
Db	1141	CCCCCGGGGCTGGCTTCGCAATTCGTCTGCTTTCACAGCTTTCAGATGCGGCTT	1200
Qy	1201	GTCCACAGCTCTTGTGTTCATCTCTCTGTGGGAGACACATCTCTCAGCTTCAA	1260
Db	1201	GTCCACAGCTCTTGTGTTCATCTCTCTGTGGGAGACACATCTCTCAGCTTCAA	1260
Qy	1261	GAGTTAGGAGAGACTCAAGTTACACTTCTCTCTCGGGTTGGAAGAAATGTTGATGC	1320

Db	1261	GAGTTAGGCGAGAGCTCAAGTTACACTTCTCTCTCTGGGGTTGGAAGAAATTGATGC	1320
Qy	1321	CAGAGGGGTGAGATTGCTGCGTCATATGAGAGCTCTCTGGAGCAAGCTCAGATGAAAA	1380
Db	1321	CAGAGGGGTGAGATTGCTGCGTCATATGAGAGCTCTCTGGAGCAAGCTCAGATGAAAA	1380
Qy	1381	GAACAAGAGGCCAGATGAGAAAGGTCTCTCTCTCTCTGGCATTAACCCAGCTTGGTT	1440
Db	1381	GAACAAGAGGCCAGATGAGAAAGGTCTCTCTCTCTCTGGCATTAACCCAGCTTGGTT	1440
Qy	1441	TGGGGGAGAGCTGGGAGAACTTCTCTCCAGAGCTGGCACTTTACGCTTGGTTAGCT	1500
Db	1441	TGGGGGAGAGCTGGGAGAACTTCTCTCCAGAGCTGGCACTTTACGCTTGGTTAGCT	1500
Qy	1501	GCTCTTGACCCCTCCACCCCGAGACAACAAGTTGGCCCGACGCTGGGCTGAC	1560
Db	1501	GCTCTTGACCCCTCCACCCCGAGACAACAAGTTGGCCCGACGCTGGGCTGAC	1560
Qy	1561	ATTGAGCCAGTGGAGCTCTGTGTCTGAAAGGGGCGTGGSCACAACCTCTTAACAACGCCCA	1620
Db	1561	ATTGAGCCAGTGGAGCTCTGTGTCTGAAAGGGGCGTGGSCACAACCTCTTAACAACGCCCA	1620
Qy	1621	CCACTTAGACACGCGCCCACTCTGACACGCGAGCTCTAGAGCTCTCTCTAGTCCCTCC	1680
Db	1621	CCACTTAGACACGCGCCCACTCTGACACGCGAGCTCTAGAGCTCTCTCTAGTCCCTCC	1680
Qy	1681	GCCCGACAGTTGTGCTTTTGTGTGGTTGGAGCTGTTTCTGTACATGATATGATGAA	1740
Db	1681	GCCCGACAGTTGTGCTTTTGTGTGGTTGGAGCTGTTTCTGTACATGATATGATGAA	1740
Qy	1741	ATGGAATCATTTGTACTGTATAAAGCTTAGTAATCCCTCTTGGGCAAGGCCCTCACCCAG	1800
Db	1741	ATGGAATCATTTGTACTGTATAAAGCTTAGTAATCCCTCTTGGGCAAGGCCCTCACCCAG	1800
Qy	1801	TTTCAGATCCAAGGCTCCACCCGGGAGCGCTTCTCTCTCTGCTGCCAACAAGGATTTCCG	1860
Db	1801	TTTCAGATCCAAGGCTCCACCCGGGAGCGCTTCTCTCTCTGCTGCCAACAAGGATTTCCG	1860
Qy	1861	TGGCCGTGTTTGGAGCTAGACATTTGACTCTCGGCATTGAGCTTCACAGGTTTACAGCAATT	1920
Db	1861	TGGCCGTGTTTGGAGCTAGACATTTGACTCTCGGCATTGAGCTTCACAGGTTTACAGCAATT	1920
Qy	1921	GCACAAGCTGGGGTGGGCAAGGCCAGAGCTGTTTTTTTATATGCTCCATTTCACAGAG	1980
Db	1921	GCACAAGCTGGGGTGGGCAAGGCCAGAGCTGTTTTTTTATATGCTCCATTTCACAGAG	1980
Qy	1981	GATTAACCAAGAACTCGAGGGGAGACACATGAGCAACAGGCCCACTTTGTCCCTAGC	2040
Db	1981	GATTAACCAAGAACTCGAGGGGAGACACATGAGCAACAGGCCCACTTTGTCCCTAGC	2040
Qy	2041	AAATTAGGGGTACAGTCCACTTAGAACCAAGGCTGCTCTATCTGTGCTGTTCTCTAAG	2100
Db	2041	AAATTAGGGGTACAGTCCACTTAGAACCAAGGCTGCTCTCTATCTGTGCTGTTCTCTAAG	2100
Qy	2101	CATTTTATTAAGACCTACTGAGGTGCTGGTTACTGTGTCTCTAGGAAACAAAGAGGTCC	2160
Db	2101	CATTTTATTAAGACCTACTGAGGTGCTGGTTACTGTGTCTCTAGGAAACAAAGAGGTCC	2160
Qy	2161	CCAGTCTTGAGCTCTGCGCCGCCCTGTGCTGCCCCACACCTTCTTGACAACAAGCGGTGGG	2220
Db	2161	CCAGTCTTGAGCTCTGCGCGCCGCCCTGTGCTGCCCCACACCTTCTTGACAACAAGCGGTGGG	2220
Qy	2221	GAGGCGGGAGAGAGAGCTGGGACCCAGAACTGAGCTGGGAGGGATCCGACAGAAAAGC	2280
Db	2221	GAGGCGGGAGAGAGAGCTGGGACCCAGAACTGAGCTGGGAGGGATCCGACAGAAAAGC	2280
Qy	2281	TCAAGGCGGGTCTTCTCTTGTGCCCCGGGATTTGGGCTATGCTGGGTTACCAATATATCTC	2340
Db	2281	TCAAGGCGGGTCTTCTCTTGTGCCCCGGGATTTGGGCTATGCTGGGTTACCAATATATCTC	2340
Qy	2341	AGGCATGTGTGGTTTGAACCCATTAACCAAGGCCCTTGTATCAGCTTTAACAAAGTA	2400
Db	2341	AGGCATGTGTGGTTTGAACCCATTAACCAAGGCCCTTGTATCAGCTTTAACAAAGTA	2400

```

QY 2401 TATTGTGATTTTATCTCTCTAAACATATTGAAGTTTAAAGGACCTTAGT 2460
Db 2401 TATTGTGATTTTATCTCTCTAAACATATTGAAGTTTAAAGGACCTTAGT 2460
QY 2461 GATCTTATTTGGGCTTTCTTGAAGTTTCAAGAGGTAAGTAATCTTCTCCAGGTTCAC 2520
Db 2461 GATCTTATTTGGGCTTTCTTGAAGTTTCAAGAGGTAAGTAATCTTCTCCAGGTTCAC 2520
QY 2521 AGCAAGTCTGGGTGGGCAAGCAAGCTAGCGCTGGGCATTCAGTACATACAGATGT 2580
Db 2521 AGCAAGTCTGGGTGGGCAAGCAAGCTAGCGCTGGGCATTCAGTACATACAGATGT 2580
QY 2581 GCTCCCTCTCTTGAAGTGGGCTTGGGCTTCAAGGCTTGGGACATCTTGGCTCA 2640
Db 2581 GCTCCCTCTCTTGAAGTGGGCTTGGGCTTCAAGGCTTGGGACATCTTGGCTCA 2640
QY 2641 ACCCTCTCCCTAGATCAAGTCTGTGAGGCTCCCTGTAGATATTGTGTACACAGTCCCATG 2700
Db 2641 ACCCTCTCCCTAGATCAAGTCTGTGAGGCTCCCTGTAGATATTGTGTACACAGTCCCATG 2700
QY 2701 TATATCAAGTACACACAGATGTACACACAGATGTACACATGCTTCAGCCCTCAGCTCTGC 2760
Db 2701 TATATCAAGTACACACAGATGTACACACAGATGTACACATGCTTCAGCCCTCAGCTCTGC 2760
QY 2761 ATACTGTGACCTGACACCCCAAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2820
Db 2761 ATACTGTGACCTGACACCCCAAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 2820
QY 2821 CCAACCTGCTCTGTCCCTTCCCAACCACTGCTGAGCTTCTGAGCAGACAGGTA 2880
Db 2821 CCAACCTGCTCTGTCCCTTCCCAACCACTGCTGAGCTTCTGAGCAGACAGGTA 2880
QY 2881 CTTGGCTGACCGGTGTGTGCTTGTACACCCAGGCAAGCCCGCAATGATC 2940
Db 2881 CTTGGCTGACCGGTGTGTGCTTGTACACCCAGGCAAGCCCGCAATGATC 2940
QY 2941 TCCGTGACACTATCAATTAAGTGGTGTGTACAAAAAAGGTTGTACAAAAAAGGTTGTACAAAAA 3000
Db 2941 TCCGTGACACTATCAATTAAGTGGTGTGTACAAAAAAGGTTGTACAAAAAAGGTTGTACAAAAA 3000
QY 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
Db 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
QY 3061 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3082
Db 3061 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3082

```

## RESULT 2

```

US-09-918-715-197
; Sequence 197, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 197
; LENGTH: 2973
; TYPE: DNA

```

ORGANISM: Homo sapiens  
US-09-918-715-197

Query Match 94.1%; Score 2900; DB 13; Length 2973;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 26 GCGCGGCCAGCCCGGGGCTCCGAGCAGCGCAGGGGAGATCCCGGCGAGTGACCCGG 85
Db 23 GCGCGGCCAGCCCGGGGCTCCGAGCAGCGCAGGGGAGATCCCGGCGAGTGACCCGG 82
QY 86 AGCCACACAGACTCTGGAGAGCTCGGCGCTGGAGAGCAGCAGAGCTCCCGAGCTCC 145
Db 83 AGCCACACAGACTCTGGAGAGCTCGGCGCTGGAGAGCAGCAGAGCTCCCGAGCTCC 142
QY 146 CGGCGCTTCCAGGAGCTCTCTGAGCCGTGCAAGAGCCGGCCCGCATTCAGACCCC 205
Db 143 CGGCGCTTCCAGGAGCTCTCTGAGCCGTGCAAGAGCCGGCCCGCATTCAGACCCC 202
QY 206 GAGGCATGATGAAACTTTGTCCAGCGGGAATGCAAGCTCAAGTCCCGCCAAACT 265
Db 203 GAGGCATGATGAAACTTTGTCCAGCGGGAATGCAAGCTCAAGTCCCGCCAAACT 262
QY 266 CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
Db 263 CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 322
QY 326 TCTCAATGCGCTTTGAGGACCAATGACACCCACATCAGAGACTTCCACCGTAAG 385
Db 323 TCTCAATGCGCTTTGAGGACCAATGACACCCACATCAGAGACTTCCACCGTAAG 382
QY 386 TATTAACATCCCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 445
Db 383 TATTAACATCCCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
QY 446 CTTTCCCGCATGCGAGGCTGCTCATCTCATCAAGGGGATGTTTATCTGTGTGTCA 505
Db 443 CTTTCCCGCATGCGAGGCTGCTCATCTCATCAAGGGGATGTTTATCTGTGTGTCA 502
QY 506 GCTGTGATACCGGAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 565
Db 503 GCTGTGATACCGGAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 562
QY 566 TCAAGTCTGCTTAAGAAACAAGCAAGAGGCGGAGCTGCCATGATGATGATGAT 625
Db 563 TCAAGTCTGCTTAAGAAACAAGCAAGAGGCGGAGCTGCCATGATGATGATGATGAT 622
QY 626 GCAACAGAAAGCAACGAGGCTGTGCGGCAAGTGTCCCAACGAGGCGAGCTGC 685
Db 623 GCAACAGAAAGCAACGAGGCTGTGCGGCAAGTGTCCCAACGAGGCGAGCTGC 682
QY 686 TGTGTGCGGGGAGAGAACTCCGCTACTTGAAGTGTGCGGCAAGAAACAACAAG 745
Db 683 TGTGTGCGGGGAGAGAACTCCGCTACTTGAAGTGTGCGGCAAGAAACAACAAG 742
QY 746 TGAACGAGATGTTTAACTGTCTTTCAGCATGCGCAAGCTGCCACAGATGAGCCCG 805
Db 743 TGAACGAGATGTTTAACTGTCTTTCAGCATGCGCAAGCTGCCACAGATGAGCCCG 802
QY 806 CCTGCAATCGCAAGATCTCCGTGAGTACGCTTCCAGCCAGGAGCTTCTGCA 865
Db 803 CCTGCAATCGCAAGATCTCCGTGAGTACGCTTCCAGCCAGGAGCTTCTGCA 862
QY 866 TGCCTCGCTCAAGAGATGAGCGCTATGAGATGATGATGATGATGATGATGATGATGAT 925
Db 863 TGCCTCGCTCAAGAGATGAGCGCTATGAGATGATGATGATGATGATGATGATGATGAT 922
QY 926 GCTTCAACAGTGAAGCTTAAGTACATCAAGGCAAGGCTTTCGGGAAAGCCAGGCGGTG 985
Db 923 GCTTCAACAGTGAAGCTTAAGTACATCAAGGCAAGGCTTTCGGGAAAGCCAGGCGGTG 982
QY 986 AGAGGACCAAGTGCACCATCAGTGAAGGAGATGCTGGGGGCTTGGCCAGTGC 1045

```

Db 983 AGAGGAGCAAGTGCACATCCAGTGAAGGAGATGCTGGGGCGGGGCTTGGCCAGTGC 1042  
Qy 1046 TTCAAGGAGGTGGCCCCAGATGCCCATGTGCGCATCTCCCAACGAGGCCCGGACCA 1105  
Db 1043 TTCAAGGAGGTGGCCCCAGATGCCCATGTGCGCATCTCCCAACGAGGCCCGGACCA 1102  
Qy 1106 GTCTTGTTCACAGACTTAAAGGACGAGCTGAGGGCCCCGGGGGCTGAGCTCCGCAAT 1165  
Db 1103 GTCTTGTTCACAGACTTAAAGGACGAGCTGAGGGCCCCGGGGGCTGAGCTCCGCAAT 1162  
Qy 1166 TCGTGTGCTTCTCAGACTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTTAT 1225  
Db 1163 TCGTGTGCTTCTCAGACTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTTAT 1222  
Qy 1226 GTCTGTGTGAGGACACATCTGTGAGCTCAAGAGTTAGGACAGACTCAAGTTACA 1285  
Db 1223 GTCTGTGTGAGGACACATCTGTGAGCTCAAGAGTTAGGACAGACTCAAGTTACA 1282  
Qy 1286 CTTTCTCTCTCGGGGGTTGGAAGAAATGTATGTCAGAGGGGTGAGGATTGCTGCTCA 1345  
Db 1283 CTTTCTCTCTCGGGGGTTGGAAGAAATGTATGTCAGAGGGGTGAGGATTGCTGCTCA 1342  
Qy 1346 TATGAGGCTCTCTGAGGACAGGCTCAGATGAAAGACACAGAGGCTCAGATGAGAA 1405  
Db 1343 TATGAGGCTCTCTGAGGACAGGCTCAGATGAAAGACACAGAGGCTCAGATGAGAA 1402  
Qy 1406 GTCTCTCTCTCTGAGGACACCCAGCTTGTGTGAGTGGAGTGGAGAACTTCTC 1465  
Db 1403 GTCTCTCTCTCTGAGGACACCCAGCTTGTGTGAGTGGAGTGGAGAACTTCTC 1462  
Qy 1466 TCCGAGCCCTGGAACCTTAAAGCTGTGTTAGCTGCTGCTGACCTCCACCCCTCA 1525  
Db 1463 TCCGAGCCCTGGAACCTTAAAGCTGTGTTAGCTGCTGCTGACCTCCACCCCTCA 1522  
Qy 1526 GCAACACACAGAGTTGGCCCCAGCTGAGCTGAGATGAGGCACTGTGTCTG 1585  
Db 1523 GCAACACACAGAGTTGGCCCCAGCTGAGCTGAGATGAGGCACTGTGTCTG 1582  
Qy 1586 AAGGGGGGCTGGCCACACTCTCTAGACAGGCCACACTTGAACACCCACCTCTC 1645  
Db 1583 AAGGGGGGCTGGCCACACTCTCTAGACAGGCCACACTTGAACACCCACCTCTC 1642  
Qy 1646 ACCGGGCTCTAGGCTCTCTCTAGGCTCTCTGCGCCGACAGTTGCTTGTGTG 1705  
Db 1643 ACCGGGCTCTAGGCTCTCTCTAGGCTCTCTGCGCCGACAGTTGCTTGTGTG 1702  
Qy 1706 TTGAGAGTGTTCGTGATGATAGTAAATGAAATGAAATCAATTGACTGTAAAGC 1765  
Db 1703 TTGAGAGTGTTCGTGATGATAGTAAATGAAATGAAATCAATTGACTGTAAAGC 1762  
Qy 1766 CTAGTACTCTCTCTCTGAGGCTCTCAACCCAGTTCAAGATCCAGGCTCCACCGG 1825  
Db 1763 CTAGTACTCTCTCTCTGAGGCTCTCAACCCAGTTCAAGATCCAGGCTCCACCGG 1822  
Qy 1826 AAGGCT 1885  
Db 1823 AAGGCT 1882  
Qy 1886 CTTCCGCAATTAGCTCAAGGTTTACAGACAAATGACAAAGCTGTGGGTGGGCCCA 1945  
Db 1883 CTTCCGCAATTAGCTCAAGGTTTACAGACAAATGACAAAGCTGTGGGTGGGCCCA 1942  
Qy 1946 GGAAGTCTTTTATATGCTCCATTTCAAGAGATACACCGAGACTCGAGGGGAC 2005  
Db 1943 GGAAGTCTTTTATATGCTCCATTTCAAGAGATACACCGAGACTCGAGGGGAC 2002  
Qy 2006 AAGATGAGACAGAGGCCCACTTTTCCCTAGCAAAATTCAGGGGTACAGCTCCACT 2065  
Db 2003 AAGATGAGACAGAGGCCCACTTTTCCCTAGCAAAATTCAGGGGTACAGCTCCACT 2062  
Qy 2066 AACCAAGGCT 2125  
Db 2063 AACCAAGGCT 2122

Qy 2126 TGGGTTCATGTGTCTTGAAGAAACAGAGAGGTCCCAAGTCTTGGGCTTGGCCGCTT 2185  
Db 2123 TGGGTTCATGTGTCTTGAAGAAACAGAGAGGTCCCAAGTCTTGGGCTTGGCCGCTT 2182  
Qy 2186 GCTGCCACACACCTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGAGCTGGAGC 2245  
Db 2183 GCTGCCACACACCTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGAGCTGGAGC 2242  
Qy 2246 CAGAACTGAGCTTGGAGGAGTCCGACAGAAAGCTCAGGGCGGGTCTTCTCTTGTGCC 2305  
Db 2243 CAGAACTGAGCTTGGAGGAGTCCGACAGAAAGCTCAGGGCGGGTCTTCTCTTGTGCC 2302  
Qy 2306 CGGATTTGGGCTATGCTGGGAGACACATGATCTCAGGCAATGGGGTTTGAACCAT 2365  
Db 2303 CGGATTTGGGCTATGCTGGGAGTCCGACAGAAAGCTCAGGGCGGGTCTTCTCTTGTGCC 2362  
Qy 2366 AACCAAGGCTCTTGTATCATGCTCTTAAACAGATATTTTGAATTTTAACTCTGAAA 2425  
Db 2363 AACCAAGGCTCTTGTATCATGCTCTTAAACAGATATTTTGAATTTTAACTCTGAAA 2422  
Qy 2426 CATATTTGAATTTTGAAGGCTTAAAGAACTTAAATGATCTTATTTGGGCTTTCTGAG 2485  
Db 2423 CATATTTGAATTTTGAAGGCTTAAAGAACTTAAATGATCTTATTTGGGCTTTCTGAG 2482  
Qy 2486 TTCAAGAGGGGTAGTAACTTCTCCAGGTGACACAGAGTCTGTGGGTGGCAGAAACA 2545  
Db 2483 TTCAAGAGGGGTAGTAACTTCTCCAGGTGACACAGAGTCTGTGGGTGGCAGAAACA 2542  
Qy 2546 AGTAGAGCTTGGGATTCAGTATACACAGATGCTCTCTCTTGAATGCTTGGCCCC 2605  
Db 2543 AGTAGAGCTTGGGATTCAGTATACACAGATGCTCTCTCTTGAATGCTTGGCCCC 2602  
Qy 2606 TGGGCTCTTCAAGGCTTGGGACATTTTGTCTTCAACCTTCCCTAGATGATGCTGGA 2665  
Db 2603 TGGGCTCTTCAAGGCTTGGGACATTTTGTCTTCAACCTTCCCTAGATGATGCTGGA 2662  
Qy 2666 GGGTCCCTGATATTTGTGTGACACAGTCCCATGATATTAACAAGTACACAGATGAT 2725  
Db 2663 GGGTCCCTGATATTTGTGTGACACAGTCCCATGATATTAACAAGTACACAGATGAT 2722  
Qy 2726 AACAGATGATACAGATGCTCCAGGCCAGCTCTGATACCTGACCTGACCCGACCTT 2785  
Db 2723 AACAGATGATACAGATGCTCCAGGCCAGCTCTGATACCTGACCTGACCCGACCTT 2782  
Qy 2786 GGGCTCTGCTGCTGTGTCTCAAGAGAGCTCCAACTGCTGTGCTTCTCTCTCTCT 2845  
Db 2783 GGGCTCTGCTGCTGTGTCTCAAGAGAGCTCCAACTGCTGTGCTTCTCTCTCTCTCT 2842  
Qy 2846 CACCCAGCTGCTGAGCTCTTGAAGACACAGTACCTTGTGACACGGGTGTGGGCC 2905  
Db 2843 CACCCAGCTGCTGAGCTCTTGAAGACACAGTACCTTGTGACACGGGTGTGGGCC 2902  
Qy 2906 GCTTCAACCCAGGACAGCCCCCGCACCATGATCTCGGTGATCACTATCAATAAAGT 2965  
Db 2903 GCTTCAACCCAGGACAGCCCCCGCACCATGATCTCGGTGATCACTATCAATAAAGT 2962  
Qy 2966 GGTGTGTACA 2976  
Db 2963 GGTGTGTACA 2973

RESULT 3  
US-09-918-715-174  
; Sequence 174, Application US/09918715  
; Publication No. US2003017157A1  
; GENERAL INFORMATION:  
; APPLICANT: Brad St. Croix  
; APPLICANT: Bert Vogelstein  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107.00134  
; CURRENT APPLICATION NUMBER: US/09/918.715

CURRENT FILING DATE: 2001-08-01  
PRIOR APPLICATION NUMBER: 60/222,599  
PRIOR FILING DATE: 2000-08-02  
PRIOR APPLICATION NUMBER: 60/224,360  
PRIOR FILING DATE: 2000-08-11  
PRIOR APPLICATION NUMBER: 60/282,850  
PRIOR FILING DATE: 2000-04-11  
NUMBER OF SEQ ID NOS: 358  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 174  
LENGTH: 2832  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-918-715-174

Query Match 89.5%; Score 2157; DB 13; Length 2832;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

199 CAGCCCGAGCCATGATGAAGACTTGTCCAGCGGGAACTGCAAGCTCAGTGTGCGCGC 258  
25 CAGCCCGAGCCATGATGAAGACTTGTCCAGCGGGAACTGCAAGCTCAGTGTGCGCGC 84  
259 AAAAACTCATACCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318  
85 AAAAACTCATACCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 144  
319 TCTGCTCTCTCAATGCGCGCTTTGAGGACCAAGTACACACCACCATTGAGAGCTTCCAC 378  
145 TCTGCTCTCTCAATGCGCGCTTTGAGGACCAAGTACACACCACCATTGAGAGCTTCCAC 204  
379 CGTAGGATTAACAATCCGCGGCAATGTAACAAGCTGCAATCCCTGATTAACCTTGGC 438  
205 CGTAGGATTAACAATCCGCGGCAATGTAACAAGCTGCAATCCCTGATTAACCTTGGC 264  
439 AACCAACCTTCCCGGCGATGCGAGCTGCTCAATCCCTCAAGGGAGTCTTCACTCG 498  
265 AACCAACCTTCCCGGCGATGCGAGCTGCTCAATCCCTCAAGGGAGTCTTCACTCG 324  
499 GTGTTCAAGCTGATTAACCGGAGTCTTCAATGAGTCAAGCGCTTCAAGAGCATC 558  
325 GTGTTCAAGCTGATTAACCGGAGTCTTCAATGAGTCAAGCGCTTCAAGAGCATC 384  
559 CTGAGAGTCAAGTCTGCTGCTGAAGAACAAGACCAAGAGGCGGAGAGTCCATGCTC 618  
385 CTGAGAGTCAAGTCTGCTGCTGAAGAACAAGACCAAGAGGCGGAGAGTCCATGCTC 444  
619 ATCTGTGCAACAAGAAAGCAACGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678  
445 ATCTGTGCAACAAGAAAGCAACGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504  
679 GAGTGTGCTGTGCTGCGGCGAGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738  
505 GAGTGTGCTGTGCTGCGGCGAGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564  
739 ACCAAGTGAAGAGATGTTTACGTGCTTTCAAGCATGCGCAAGCTTCCACAGAGATG 798  
565 ACCAAGTGAAGAGATGTTTACGTGCTTTCAAGCATGCGCAAGCTTCCACAGAGATG 624  
799 AGCCCGCCCTGCAATGCAAGATCTCGTGAAGTACAGGTGAGCTTCCACCCAGGCGC 858  
625 AGCCCGCCCTGCAATGCAAGATCTCGTGAAGTACAGGTGAGCTTCCACCCAGGCGC 684  
859 TTCTGATGCGCGCTGCAAGAGATGAGCGCTTATGAGCATGCTGCTGCTGCTGCTGCTG 918  
685 TTCTGATGCGCGCTGCAAGAGATGAGCGCTTATGAGCATGCTGCTGCTGCTGCTGCTG 744  
919 CGCCCGAGCTTCAAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 978  
745 CGCCCGAGCTTCAAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 804  
979 GCCCGTGAAGAGGACCAAGTGAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1038

805 GCCCGTGAAGAGGACCAAGTGAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 864  
1039 CAGTGTCTTCAAGGAGGAGTGGGCGGAGTGGGCGGAGTGGGCGGAGTGGGCGGAGTGGGCGG 1098  
865 CAGTGTCTTCAAGGAGGAGTGGGCGGAGTGGGCGGAGTGGGCGGAGTGGGCGGAGTGGGCGG 924  
1099 GGCAGAGTCTTGTTCACAGACTTAAAGCAACCAAGTGAAGGCGGCGGAGTGGGCGGAGTGGGCGG 1158  
925 GGCAGAGTCTTGTTCACAGACTTAAAGCAACCAAGTGAAGGCGGCGGAGTGGGCGGAGTGGGCGG 984  
1159 GGCAGAGTCTTGTTCACAGACTTAAAGCAACCAAGTGAAGGCGGCGGAGTGGGCGGAGTGGGCGG 1218  
985 GGCAGAGTCTTGTTCACAGACTTAAAGCAACCAAGTGAAGGCGGCGGAGTGGGCGGAGTGGGCGG 1044  
1219 GTCATATCTCTCTGCTGAGGAGGAGACATCTTGAAGTCAAGGTTAGGAGAGTCTCA 1278  
1045 GTCATATCTCTCTGCTGAGGAGGAGACATCTTGAAGTCAAGGTTAGGAGAGTCTCA 1104  
1279 AGTTACACCTTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1338  
1105 AGTTACACCTTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1164  
1339 TGCGTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1398  
1165 TGCGTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1224  
1399 GAGAAAGTCT 1458  
1225 GAGAAAGTCT 1284  
1459 ACTTCT 1518  
1285 ACTTCT 1344  
1519 ACCCGAGACACACACAGAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1578  
1345 ACCCGAGACACACACAGAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1404  
1579 GTGTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1638  
1405 GTGTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1464  
1639 CTTCTGAGCGGCTTCT 1698  
1465 CTTCTGAGCGGCTTCT 1524  
1699 GTTGTGAGTGAAGTCT 1758  
1525 GTTGTGAGTGAAGTCT 1584  
1759 TAAAGGCTTGAAGTCT 1818  
1585 TAAAGGCTTGAAGTCT 1644  
1819 ACCCGAGACGCTTCT 1878  
1645 ACCCGAGACGCTTCT 1704  
1879 ACATTGAAGTCT 1938  
1705 ACATTGAAGTCT 1764  
1939 CAGGCGAGAGTCT 1998  
1765 CAGGCGAGAGTCT 1824  
1999 AGGGGAGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2058  
1825 AGGGGAGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1884  
2059 CACCTAGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2118  
1885 CACCTAGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1944

QY 2119 TGGGTCTGGGTCTCACTGTCTCTAGAAACCAAGGGTCCCAAGTCTTGGCTCTCC 2178  
DB 1945 TGGGTCTGGGTCTCACTGTCTCTAGAAACCAAGGGTCCCAAGTCTTGGCTCTCC 2004  
QY 2179 CGCCCTCTGCTGCCCCACCACTTCTGCAACAAGCGGGGAGGCGGGAGAGCAGC 2238  
DB 2005 CGCCCTCTGCTGCCCCACCACTTCTGCAACAAGCGGGGAGGCGGGAGAGCAGC 2064  
QY 2239 TGGGACCCAGAACTGAGCTGGGAGGGAATCCGACAGAAAAGCTCAAGGCGGGTCTTCC 2298  
DB 2065 TGGGACCCAGAACTGAGCTGGGAGGGAATCCGACAGAAAAGCTCAAGGCGGGTCTTCC 2124  
QY 2299 TTGTGCCCCGGGATTTGGGCTATGCTGGGTAACAATGTAATCTTATTTTAAATCT 2358  
DB 2125 TTGTGCCCCGGGATTTGGGCTATGCTGGGTAACAATGTAATCTTATTTTAAATCT 2184  
QY 2359 ACCCATAAACCAAGGCGCTTGTATGATGATCTTAAAGTATTTTGTATTTTAAATCT 2418  
DB 2185 ACCCATAAACCAAGGCGCTTGTATGATGATCTTAAAGTATTTTGTATTTTAAATCT 2244  
QY 2419 CTCTAATATTTAATTTTAAAGTCTTAAAGGCGCTTAAAGAACTTAAAGTATTTTAAATCT 2478  
DB 2245 CTCTAATATTTAATTTTAAAGTCTTAAAGGCGCTTAAAGAACTTAAAGTATTTTAAATCT 2304  
QY 2479 TCTGAGGTTCAAGAGAGGTAAGTAATCTTCTCAAGTCAACAAGCAAGTCTTGGGTGGC 2538  
DB 2305 TCTGAGGTTCAAGAGAGGTAAGTAATCTTCTCAAGTCAACAAGCAAGTCTTGGGTGGC 2364  
QY 2539 AGAAGCAAGCTAAGGCTGGGCACTTCAAGTCAATACAGATGATGCTTCTTATGATCT 2598  
DB 2365 AGAAGCAAGCTAAGGCTGGGCACTTCAAGTCAATACAGATGATGCTTCTTATGATCT 2424  
QY 2599 TGGCCCTTGGGCTTCAAGGCTTGGGCACTTGTCTCAACCTCTCTCTAGATCAG 2658  
DB 2425 TGGCCCTTGGGCTTCAAGGCTTGGGCACTTGTCTCAACCTCTCTCTAGATCAG 2484  
QY 2659 TCTGAGGTTCTCTGTAATTTTGTATGTAACCAATGCTTATTTATTAATCAACA 2718  
DB 2485 TCTGAGGTTCTCTGTAATTTTGTATGTAACCAATGCTTATTTATTAATCAACA 2544  
QY 2719 GATGTAACAAGATGTAACATGCTCCAGGCGCTGATCTGATCTGATCTGATCTGATCT 2778  
DB 2545 GATGTAACAAGATGTAACATGCTCCAGGCGCTGATCTGATCTGATCTGATCTGATCT 2604  
QY 2779 CAGCTTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2838  
DB 2605 CAGCTTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2664  
QY 2839 CCTTCCCAACCACTGCTGAGCTTCTGAGCAACAGTACCTTGGCTGACCGGTCT 2898  
DB 2665 CCTTCCCAACCACTGCTGAGCTTCTGAGCAACAGTACCTTGGCTGACCGGTCT 2724  
QY 2899 GTGGCCCGCTCTCAACCAAGGCAAGGCGGCAACATGATCTCGGTGATCACTATCAAT 2958  
DB 2725 GTGGCCCGCTCTCAACCAAGGCAAGGCGGCAACATGATCTCGGTGATCACTATCAAT 2784  
QY 2959 AAAAATGGGTTTGTACAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 3006  
DB 2785 AAAAATGGGTTTGTACAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 2832

RESULT 4  
US-10-185-035-4

; Sequence 4, Application US/10185035  
; Publication No. US20040005706A1

; GENERAL INFORMATION:  
; APPLICANT: Brett P. Monia

; APPLICANT: Kenneth W. Dobie  
; TITLE OF INVENTION: ANTISENSE MODULATION OF DEXRAS1 EXPRESSION

; FILE REFERENCE: RTS-0377  
; CURRENT APPLICATION NUMBER: US/10/185,035

; CURRENT FILING DATE: 2002-06-28

NUMBER OF SEQ ID NOS: 135  
; SEQ ID NO 4  
; LENGTH: 2832  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(837)  
US-10-185-035-4

Query Match 89.5%; Score 2757; DB 16; Length 2832;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 CAGCCCGAGGCAATGATGAGAACTTTGTCCAGCGGGAATCTGACGCTCAGTGTCCCGCC 258  
DB 25 CAGCCCGAGGCAATGATGAGAACTTTGTCCAGCGGGAATCTGACGCTCAGTGTCCCGCC 84  
QY 259 AAAAATCATACCGCATGCTGTGCTGAGTCTCTCGGGTGGGCAAGAGCTCCATCGTG 318  
DB 85 AAAAATCATATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 144  
QY 319 TCTGCTTCTCAATATGCGCTTGTAGAGACCAAGTCAACCCACATCGAGACTTCCAC 378  
DB 145 TCTGCTTCTCAATATGCGCTTGTAGAGACCAAGTCAACCCACATCGAGACTTCCAC 204  
QY 379 CGTAAAGTATCAATCTCCGCGGCAATGTAACCAAGTCAATCTTGGATATCTTGGC 438  
DB 205 CGTAAAGTATCAATCTCCGCGGCAATGTAACCAAGTCAATCTTGGATATCTTGGC 264  
QY 439 AACCAACCTTCCCGGCGCATGCGAGGCTGTCATCTCAAGGGAATGCTTATCTGCTG 498  
DB 265 AACCAACCTTCCCGGCGCATGCGAGGCTGTCATCTCAAGGGAATGCTTATCTGCTG 324  
QY 499 GTGTTCAGCTGATTAACCGGAGTCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 558  
DB 325 GTGTTCAGCTGATTAACCGGAGTCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 384  
QY 559 CTGAGGTTCAAGTCTGCTGTAAGAAACAAGCAAGAGAGGCGGAGTCCATGCTC 618  
DB 385 CTGAGGTTCAAGTCTGCTGTAAGAAACAAGCAAGAGAGGCGGAGTCCATGCTC 444  
QY 619 ATCTGTGCAACAAGAAACAAGCAAGGAGCTGTGCGGCGGAGTCCCAACCAAGAGGCT 678  
DB 445 ATCTGTGCAACAAGAAACAAGCAAGGAGCTGTGCGGCGGAGTCCCAACCAAGAGGCT 504  
QY 679 GAGTCTGCTGTGCGGCGGAGGAACTCGCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 738  
DB 505 GAGTCTGCTGTGCGGCGGAGGAACTCGCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 564  
QY 739 ACCAAGTGAACAGATTTCTACGTGCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 798  
DB 565 ACCAAGTGAACAGATTTCTACGTGCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 624  
QY 799 AGCCCGGCTGATCGCAAGTCTCGTGAAGTGAAGGCTTCAACCCAGAGGCT 858  
DB 625 AGCCCGGCTGATCGCAAGTCTCGTGAAGTGAAGGCTTCAACCCAGAGGCT 684  
QY 859 TTCTGATGCGCGCGGTCAAGAGATGAGCGCTTATGATGATGATGATGATGATGATGATGAT 918  
DB 685 TTCTGATGCGCGCGGTCAAGAGATGAGCGCTTATGATGATGATGATGATGATGATGATGAT 744  
QY 919 CGCCCGAGGTCACAGTGAACCTCAAGTCAATCAAGGCAAGTCTTCCGGAAGGCGAG 978  
DB 745 CGCCCGAGGTCACAGTGAACCTCAAGTCAATCAAGGCAAGTCTTCCGGAAGGCGAG 804  
QY 979 GCGCGTGAAGAGGACAAGTGAACCAATCAAGAGAGGAGATGCTGGGCGGGGCTTGGC 1038  
DB 805 GCGCGTGAAGAGGACAAGTGAACCAATCAAGAGAGGAGATGCTGGGCGGGGCTTGGC 864  
QY 1039 CAGTCTTCAAGGAGTGGGCGGCAAGATGCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 1098  
DB 865 CAGTCTTCAAGGAGTGGGCGGCAAGATGCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 924

1099 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1158  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 984  
Qy 1159 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1218  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1044  
Qy 1219 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1278  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1104  
Qy 1279 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1338  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1164  
Qy 1339 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1398  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1224  
Qy 1399 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1458  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1284  
Qy 1459 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1518  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1344  
Qy 1519 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1578  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1404  
Qy 1579 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1638  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1464  
Qy 1639 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1698  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1524  
Qy 1698 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1758  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1584  
Qy 1759 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1818  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1644  
Qy 1819 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1878  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1704  
Qy 1879 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1938  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1764  
Qy 1939 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1998  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1824  
Qy 1999 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2058  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1884  
Qy 2059 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2118  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 1944  
Qy 2119 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2178  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2004

Qy 2179 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2238  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2064  
Qy 2239 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2298  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2124  
Qy 2299 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2358  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2184  
Qy 2359 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2418  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2244  
Qy 2419 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2478  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2304  
Qy 2479 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2538  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2364  
Qy 2539 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2598  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2424  
Qy 2599 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2658  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2484  
Qy 2659 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2718  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2544  
Qy 2719 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2778  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2604  
Qy 2779 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2838  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2664  
Qy 2839 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2898  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2724  
Qy 2899 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2958  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2784  
Qy 2959 GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 3006  
Db GGCAGACAGTCTTGTTCACAGACCTTAGGACACAGACTGGAGGCCCGCGGCGCTGGCTTC 2832

RESULT 5  
US-09-778-963a-3  
; Sequence 3, Application US/09778963A  
; Patent No. US20020115172A1  
; GENERAL INFORMATION:  
; APPLICANT: NEULAM, Beena et al  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; FILE REFERENCE: C0001112  
; CURRENT APPLICATION NUMBER: US/09/778, 963A  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 11221







Dh 9731 ACAGCAAGTCTGTGGTGGGAGAGCAAGCTAGCGCTGGCATTCAGTACATACACGAT 9790  
Qy 2579 GTGTCCTCTCTGATGATGCTGGCCCTGGGGCCCTTACAGGCTTTGGAGACATCTTGTCT 2638  
Dh 9791 GTGTCCTCTCTGATGATGCTGGCCCTGGGGCCCTTACAGGCTTTGGAGACATCTTGTCT 9850  
Qy 2639 CAACCTCTCTGATGATGCTGGAGGCTCCCTGATATTTGTGACACATGCCCCA 2698  
Dh 9851 CAACCTCTCTGATGATGCTGGAGGCTCCCTGATATTTGTGACACATGCCCCA 9910  
Qy 2699 TGTATATACAGTACACACAGATGTACACAGATGTACACATGTCTCCAGCCACTCT 2758  
Dh 9911 TGTATATACAGTACACACAGATGTACACAGATGTACACATGTCTCCAGCCACTCT 9970  
Qy 2759 GCATACCTGACCTGACACCCAGCCTGGCCCTGCTGCTGTGTCTTCAAGCAGAG 2818  
Dh 9971 GCATACCTGACCTGACACCCAGCCTGGCCCTGCTGCTGTGTCTTCAAGCAGAG 10030  
Qy 2819 CTCACACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2878  
Dh 10031 CTCACACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10090  
Qy 2879 TACCTTGGCTGACACCGGTGTGGCCGCTCTCAACCCAGGACACGCCGCCACCATGA 2938  
Dh 10091 TACCTTGGCTGACACCGGTGTGGCCGCTCTCAACCCAGGACACGCCGCCACCATGA 10150  
Qy 2939 TCTCCGTGTACATCATATAAAGTGGTGTGTACAA 2978  
Dh 10151 TCTCCGTGTACATCATATAAAGTGGTGTGTACAA 10190

## RESULT 6

US-10-185-035-11  
; Sequence 11, Application US/10185035  
; Publication No. US20040005706A1  
; GENERAL INFORMATION:  
; APPLICANT: Brett P. Monia  
; APPLICANT: Kenneth W. Dobie  
; TITLE OF INVENTION: ANTISENSE MODULATION OF DEXRASI EXPRESSION  
; FILE REFERENCE: RFS-0377  
; CURRENT APPLICATION NUMBER: US/10/185,035  
; CURRENT FILING DATE: 2002-06-28  
; NUMBER OF SEQ ID NOS: 135  
; SEQ ID NO 11  
; LENGTH: 16000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-185-035-11

Query Match 79.5%; Score 2449; DB 16; Length 16000;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 479 CAGGGAGATGTTTATCTATCTGTGTTCAAGCTGATTAACGGAGAGCTTTCATGATAGTCA 538  
Dh 11923 CAGGGAGATGTTTATCTATCTGTGTTCAAGCTGATTAACGGAGAGCTTTCATGATAGTCA 11982  
Qy 539 AGCGCTTCAGAGCAGATCTGAGAGTCAAGTCTGCTTGAAGAACAAAGCAAGAGAGG 598  
Dh 11983 AGCGCTTCAGAGCAGATCTGAGAGTCAAGTCTGCTTGAAGAACAAAGCAAGAGAGG 12042  
Qy 599 CGGGGAGCTGGCCATGATGATCTGTGGCAACAGAACAGCAGCGGAGCTGTGGCCG 658  
Dh 12043 CGGGGAGCTGGCCATGATGATCTGTGGCAACAGAACAGCAGCGGAGCTGTGGCCG 12102  
Qy 659 AGGTGCCACACGAGGCGAGCTGTGTGTGGGAGAGAGAACTCCGCTACTTCTG 718  
Dh 12103 AGGTGCCACACGAGGCGAGCTGTGTGTGGGAGAGAGAACTCCGCTACTTCTG 12162  
Qy 719 AGGTGTGGCCAGAGAGAACCAACGTTGAGCAGATGTTCTACGTGCTTTCAGCATGG 778  
Dh 12163 AGGTGTGGCCAGAGAGAACCAACGTTGAGCAGATGTTCTACGTGCTTTCAGCATGG 12222

Qy 779 CCAAGCTGCCACAGAGATGAGCCCGCCCTGATCTCCAGATGATCTCCGTGACATGCGTGG 838  
Dh 12223 CCAAGCTGCCACAGAGATGAGCCCGCCCTGATCTCCAGATGATCTCCGTGACATGCGTGG 12282  
Qy 839 AGCGCTTCCACCCAGGCGCTTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGGCA 898  
Dh 12283 AGCGCTTCCACCCAGGCGCTTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGGCA 12342  
Qy 899 TGTGTGCGCCCTTGGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 958  
Dh 12343 TGTGTGCGCCCTTGGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 12402  
Qy 959 AGGTCTTGGGAGAGGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1018  
Dh 12403 AGGTCTTGGGAGAGGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 12462  
Qy 1019 ATGTGGGCGGCGCTTGGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1078  
Dh 12463 ATGTGGGCGGCGCTTGGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 12522  
Qy 1079 CATCTCCCAACCGAGGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1138  
Dh 12523 CATCTCCCAACCGAGGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 12582  
Qy 1139 GGCCCGCGGCGCTGGCGCTCGCACATTCGTGCTGCTCTCAAGCTTTCCTGAGTCCGC 1198  
Dh 12583 GGCCCGCGGCGCTGGCGCTCGCACATTCGTGCTGCTCTCAAGCTTTCCTGAGTCCGC 12642  
Qy 1199 TTTGTCAAGCTCTTGGTGGTTTCACTCTCTGTGGAGAGACATCTGTGACGCTC 1258  
Dh 12643 TTTGTCAAGCTCTTGGTGGTTTCACTCTCTGTGGAGAGACATCTGTGACGCTC 12702  
Qy 1259 AAGATTAGGAGAGACTCAAGTTTACCTCTCTCTCTGCGGTTTGAAGAAAGTTTAT 1318  
Dh 12703 AAGATTAGGAGAGACTCAAGTTTACCTCTCTCTCTGCGGTTTGAAGAAAGTTTAT 12762  
Qy 1319 GCCAGAGGAGGAGATTGCTGCTCATATGAGAGCTCTGAGCAAGCTCAGAGATGA 1378  
Dh 12763 GCCAGAGGAGGAGATTGCTGCTCATATGAGAGCTCTGAGCAAGCTCAGAGATGA 12822  
Qy 1379 AAGACACAGAGGCGAGATGAGAAAGTCTCTCTCTCTGAGCAATCAACCCAGCTTGG 1438  
Dh 12823 AAGACACAGAGGCGAGATGAGAAAGTCTCTCTCTCTGAGCAATCAACCCAGCTTGG 12882  
Qy 1439 TTTGGTGGCAGCTGGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGCTGTGTAG 1498  
Dh 12883 TTTGGTGGCAGCTGGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGCTGTGTAG 12942  
Qy 1499 CTGCTCTGACACCCCTCCCAACCCCGACACACACACAAAGTTGGCCCGCCAGCTGGCGCTG 1558  
Dh 12943 CTGCTCTGACACCCCTCCCAACCCCGACACACACAAAGTTGGCCCGCCAGCTGGCGCTG 13002  
Qy 1559 ACATTGAGCCAGTGAAGTCTGTGTGTAAGGGGCGTGGCCACACTCTTGAACCAAGCC 1618  
Dh 13003 ACATTGAGCCAGTGAAGTCTGTGTGTAAGGGGCGTGGCCACACTCTTGAACCAAGCC 13062  
Qy 1619 CACCACTTGAACACCGCCCACTCTGACCGCTTCTCAAGCTCTTCTTCAAGTCCCT 1678  
Dh 13063 CACCACTTGAACACCGCCCACTCTGACCGCTTCTCAAGCTCTTCTTCAAGTCCCT 13122  
Qy 1679 CGGCCGACAGTTGCTTGTGTGTGGTGGAGCTGTTTGTGTATGTATGATAG 1738  
Dh 13123 CGGCCGACAGTTGCTTGTGTGTGGTGGAGCTGTTTGTGTATGTATGATAG 13182  
Qy 1739 AAATGAAATCATTTGATCTGTAAAGCTTATGATCTCCCTCTTGGCCAGGCGCTCACCC 1798  
Dh 13183 AAATGAAATCATTTGATCTGTAAAGCTTATGATCTCCCTCTTGGCCAGGCGCTCACCC 13242  
Qy 1799 AGTTCAAGTCCAGGCGCTTCCACCGGAGAGCGCTTCTCTGTGCTCCCAACAGAGTTTC 1858  
Dh 13243 AGTTCAAGTCCAGGCGCTTCCACCGGAGAGCGCTTCTCTGTGCTCCCAACAGAGTTTC 13302  
Qy 1859 CGTGAGCTGTTTGACGTAGACATTTAGCTCCGCAATTTAGCTCCAGGTTTACAGACAA 1918

```

Db      13303  CGTGGCTGTTGACGTAGACATTGACCTCCGCAATTGAGCTCCACGGTTTACAGACAA 13362
Oy      1919    TTGCACAAGCGTGGGGTGGGAGGCGCAGAGCTGCTTTTATTAATGCTCCATTTCACAG 1978
Db      13363  TTGCACAAGCGTGGGGTGGGAGGCGCAGAGCTGCTTTTATTAATGCTCCATTTCACAG 13422
Oy      1979    AGGATACACCGAGACTCGAGAGGGGACACGATGAGACACAGGCCCTTATTCCTCCCTA 2038
Db      13423  AGGATACACCGAGACTCGAGAGGGGACACGATGAGACACAGGCCCTTATTCCTCCCTA 13482
Oy      2039  GCAATTTACAGGGTACAGGCTCCACCTAGAACAGGAGGCTGCTGCTGCTGCTGCTCA 2098
Db      13483  GCAATTTACAGGGTACAGGCTCCACCTAGAACAGGAGGCTGCTGCTGCTGCTGCTCA 13542
Oy      2099  AGCATTTTATTAAGCACTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158
Db      13543  AGCATTTTATTAAGCACTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13602
Oy      2159  CCCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2218
Db      13603  CCCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13662
Oy      2219  GGGAGCGGGGAGAGAGAGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2278
Db      13663  GGGAGCGGGGAGAGAGAGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 13722
Oy      2279  GCTCAAGGGGGGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338
Db      13723  GCTCAAGGGGGGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13782
Oy      2339  TGAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2398
Db      13783  TGAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13842
Oy      2399  TATATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2458
Db      13843  TATATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 13902
Oy      2459  GTGATCTTCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2518
Db      13903  GTGATCTTCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13962
Oy      2519  AAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2578
Db      13963  AAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14022
Oy      2579  GTGCTCCCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2638
Db      14023  GTGCTCCCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14082
Oy      2639  CAACTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2698
Db      14083  CAACTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14142
Oy      2699  TGTATTTAAGTACACAGATGTAACAGATGTAACAGATGTAACAGATGTAACAGATGTAAC 2758
Db      14143  TGTATTTAAGTACACAGATGTAACAGATGTAACAGATGTAACAGATGTAACAGATGTAAC 14202
Oy      2759  GATATCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 2818
Db      14203  GATATCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 14262
Oy      2819  CTCGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2878
Db      14263  CTCGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 14322
Oy      2879  TACCTTGGCTGCAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2938
Db      14323  TACCTTGGCTGCAACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14382
Oy      2939  TCTCCGTGTAACATCAATTAAGTGGGTTGTTCAAA 2978

```

```

Db      14383  TCTCCGTGTAACATCAATTAAGTGGGTTGTTCAAA 14422

RESULT 7
US-10-321-039-13
; Sequence 13, Application US/10321039
; Publication No. US20040014067A1
; GENERAL INFORMATION:
; APPLICANT: Lyumichev, Victor
; APPLICANT: Lukowick, Andrew
; APPLICANT: Jarvis, Nancy
; APPLICANT: Kurensky, David
; TITLE OF INVENTION: Amplification Methods and Compositions
; FILE REFERENCE: FORS-06960
; CURRENT APPLICATION NUMBER: US/10/321,039
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: 09/998,157
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/329,113
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/360,489
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 759
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (561)..(561)
; OTHER INFORMATION: n can be g or t.
US-10-321-039-13

Query Match      30.1%; Score 927; DB 16; Length 1021;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 977; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      708    GCGCTACTTGAGGGTGGGCGCAAGAGACCAAGTGAAGTGTCTAGTCT 767
Db      44     CGCTACTTCAGAGTGTGGCGCAAGAGACCAAGTGAAGTGTCTAGTCT 103
Oy      768    CTTGAGCATGCGCAAGCTGCGACAGAGATGAGCCCGCTGATGCAAGATCTCCGT 827
Db      104    CTTGAGCATGCGCAAGCTGCGACAGAGATGAGCCCGCTGATGCAAGATCTCCGT 163
Oy      828    GCAATGCGTGAAGCTTTCACCCCAAGGCTTTCGATGAGCGCGCGCAAGAGATGGA 887
Db      164    GCAATGCGTGAAGCTTTCACCCCAAGGCTTTCGATGAGCGCGCGCAAGAGATGGA 223
Oy      888    GCGCTATGAGCATGCTGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 947
Db      224    GCGCTATGAGCATGCTGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 283
Oy      948    CATCAAGGCAAGTCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1007
Db      284    CATCAAGGCAAGTCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 343
Oy      1008  GTGAGCGAGGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1067
Db      344    GTGAGCGAGGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 403
Oy      1068  CCCACTGTGCGCATCTCCCAACCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1127
Db      404    CCCACTGTGCGCATCTCCCAACCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 463
Oy      1128  ACCAGACTGAGAGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1187
Db      464    ACCAGACTGAGAGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 523
Oy      1188  CCTGAGTCCGCTTTCACAGTCTCTGAGTGTTCATCTCTGAGGAGGAGACATC 1247
Db      524    CCTGAGTCCGCTTTCACAGTCTCTGAGTGTTCATCTCTGAGGAGGAGACATC 583

```

QY 1248 TCTGACCTTCAAGAGTTAGGACAGACTCAAGTTACACTTCTCTCTGGGGTTGAA 1307  
DB 584 TCTGACCTTCAAGAGTTAGGACAGACTCAAGTTACACTTCTCTCTGGGGTTGAA 643  
QY 1308 GAATGTTGATGACAGAGGGGTGAGATTGCTGCTCATATGAGACCTCTGGACAAGC 1367  
DB 644 GAATGTTGATGACAGAGGGGTGAGATTGCTGCTCATATGAGACCTCTGGACAAGC 703  
QY 1368 CTCAGATGAAAAGACACAGAAAGCCAGATGAAAGGTCTCTCTCTGACATTAAC 1427  
DB 704 CTCAGATGAAAAGACACAGAAAGCCAGATGAAAGGTCTCTCTCTCTGACATTAAC 763  
QY 1428 ACCAGCTGTTGGTGGTGGAGCTGAGAACTTCTCTCTGACACCTGGAACCTTAAG 1487  
DB 764 ACCAGCTGTTGGTGGTGGAGCTGAGAACTTCTCTCTGACACCTGGAACCTTAAG 823  
QY 1488 CTCTGTTGAGCTCTCTGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1547  
DB 824 CTCTGTTGAGCTCTCTGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 883  
QY 1548 AGCTGCTCTGACATTTAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1607  
DB 884 AGCTGCTCTGACATTTAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943  
QY 1608 TAGACCACTGACATTTAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1667  
DB 944 TAGACCACTGACATTTAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003  
QY 1668 CCTAGTCT 1685  
DB 1004 CCTAGTCT 1021

## RESULT 8

US-10-029-386-20951  
; Sequence 20951, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 20951  
; LENGTH: 696  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL022334.1  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.6  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
; OTHER INFORMATION: EST HUMAN HIT: B1754083.1, EVALUATE 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: P10301, EVALUATE 4.00e-10  
; OTHER INFORMATION: NT HIT: g16168170, EVALUATE 0.00e+00  
US-10-029-386-20951

Query Match 20.9%; Score 645; DB 15; Length 696;  
Best Local Similarity 99.9%; Pred. No. 1.7e-299;

Matches 695; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 482 GGGATGCTTATCTCTGTTTCACTGATTAACCGGAGTCTTTCATGATGAGTCAAGC 541  
DB 1 GGGATGCTTATCTCTGTTTCACTGATTAACCGGAGTCTTTCATGATGAGTCAAGC 60  
QY 542 GCTTCAAGACAGATCTGAGGCTCAAGTCTGCTGAAGAACCAAGACCAAGAGCGG 601

DB 61 GCTTCAAGACAGATCTGAGGCTCAAGTCTGCTGAAGAACCAAGACCAAGAGCGG 120  
QY 602 CGAGCTGCTCATCTGTTGCAAGAAACACACAGGAGACTGTGCGCCAG 661  
DB 121 CGAGCTGCTCATCTGTTGCAAGAAACACACAGGAGACTGTGCGCCAG 180  
QY 662 TGCCACACCGAGGCGGAGTGTGTTGCGGCGACGAGAACTCGGCTACTTTCAGG 721  
DB 181 TGCCACACCGAGGCGGAGTGTGTTGCGGCGACGAGAACTCGGCTACTTTCAGG 240  
QY 722 TGTCGCAAGAAAGAACCAACCTGACAGATGTTCTAGTCTCTTTCAGATGCA 781  
DB 241 TGTCGCAAGAAAGAACCAACCTGACAGATGTTCTAGTCTCTTTCAGATGCA 300  
QY 782 AGCTGCCACACAGATGAGCCCGCTGCTGATGCGAAGATCTCCGTGACATGAGTACG 841  
DB 301 AGCTGCCACACAGATGAGCCCGCTGCTGATGCGAAGATCTCCGTGACATGAGTACG 360  
QY 842 CCTTCAACCCAGGCTCTGATGAGCGCGCTGCAAGAGATGAGAGCTTATGAGATG 901  
DB 361 CCTTCAACCCAGGCTCTGATGAGCGCGCTGCAAGAGATGAGAGCTTATGAGATG 420  
QY 902 TCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961  
DB 421 TCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 962 TCTTCTGGAAGGCGAGGCGCTGAGAGGCAAGTGCATCTCAGTGAAGGAGATG 1021  
DB 481 TCTTCTGGAAGGCGAGGCGCTGAGAGGCAAGTGCATCTCAGTGAAGGAGATG 540  
QY 1022 CTGAGGCGGCGCTGAGGCGAGGCTTCAAGGAGGTTGAGGCGGCTGAGGCTGAGT 1081  
DB 541 CTGAGGCGGCGCTGAGGCGAGGCTTCAAGGAGGTTGAGGCGGCTGAGGCTGAGT 600  
QY 1082 CTCCCAACGAGGCGGCGAGGCTTGTTCACAGACCTTGAAGCACCACTGAGAGC 1141  
DB 601 CTCCCAACGAGGCGGCGAGGCTTGTTCACAGACCTTGAAGCACCACTGAGAGC 660  
QY 1142 CCCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177  
DB 661 CCCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696

## RESULT 9

US-10-027-632-25123/c  
; Sequence 25123, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: Polymorphisms in the Human Genome  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25123  
; LENGTH: 753  
; TYPE: DNA

ORGANISM: Human  
US-10-027-632-25123

Query Match 18.6%; Score 574; DB 13; Length 753;  
Best Local Similarity 99.6%; Pred. No. 2.3e-265;  
Matches 724; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

718 GAGGTGTGGGCGCAAGAAACCAACCGTGAAGAGATGTTTACGTGCTTTCAGCATG 777  
731 GAGGTGTGGGCGCAAGAAACCAACCGTGAAGAGATGTTTACGTGCTTTCAGCATG 672  
778 GCCAAGCTGCCACACGAGATGAGCCCGCCCTGCAATGCAAGATCTCCGTGAGTACG 837  
671 GCCAAGCTGCCACACGAGATGAGCCCGCCCTGCAATGCAAGATCTCCGTGAGTACG 612  
838 GACGCTTCCACACCGAGGCTTCTGCAATGCGCCGCTGAAGAGATGAGCCCTATG 897  
611 GACGCTTCCACACCGAGGCTTCTGCAATGCGCCGCTGAAGAGATGAGCCCTATG 552  
898 ATGATCTGCGCCCTTGCAGCCCGCCCGCCAGCGTCAACAGTACCTGAAGTACCAAGGCC 957  
551 ATGATCTGCGCCCTTGCAGCCCGCCCGCCAGCGTCAACAGTACCTGAAGTACCAAGGCC 492  
958 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 1017  
491 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 432  
1018 GATGCTGGGGGGGGGCTTGGCCAGTGCCTTCAAGGAGTGGGCGCCCAAGTGCCTGTC 1077  
431 GATGCTGGGGGGGGGCTTGGCCAGTGCCTTCAAGGAGTGGGCGCCCAAGTGCCTGTC 372  
1078 GCATCTCCCAACCGAGGCGCCCGCCAGCGTCTTGTTCACAGACTTAAAGCAACAGCTGG 1137  
371 GCATCTCCCAACCGAGGCGCCCGCCAGCGTCTTGTTCACAGACTTAAAGCAACAGCTGG 312  
1138 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 1197  
311 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 252  
1198 CTGTGCCAAGCTCTTGTGTTTCAATCTCTCTGTGGAGAGACAACTCTCTGAGCT 1257  
251 CTGTGCCAAGCTCTTGTGTTTCAATCTCTCTGTGGAGAGACAACTCTCTGAGCT 192  
1258 CAAGATTAGGAGAGACTCAAGTTCACCTCTCTCTGAGGTTGGAGAAATGTTGA 1317  
191 CAAGATTAGGAGAGACTCAAGTTCACCTCTCTCTGAGGTTGGAGAAATGTTGA 132  
1318 TGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACCTCTGAGCAAGCTCAGAGTGA 1377  
131 TGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACCTCTGAGCAAGCTCAGAGTGA 72  
1378 AAAGGACACAGAAAGGCGCAGATGAGAAAGTCTCTCTCTGAGGATTAACCCGAGCTTG 1437  
71 AAAGGACACAGAAAGGCGCAGATGAGAAAGTCTCTCTCTGAGGATTAACCCGAGCTTG 12  
1438 GTTTGGG 1444  
11 GTTTGGG 5

RESULT 10  
US-10-027-632-25124/c  
; Sequence 25124, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 25124  
; LENGTH: 753  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-25124

Query Match 18.6%; Score 574; DB 13; Length 753;  
Best Local Similarity 99.6%; Pred. No. 2.3e-265;  
Matches 724; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

718 GAGGTGTGGGCGCAAGAAACCAACCGTGAAGAGATGTTTACGTGCTTTCAGCATG 777  
731 GAGGTGTGGGCGCAAGAAACCAACCGTGAAGAGATGTTTACGTGCTTTCAGCATG 672  
778 GCCAAGCTGCCACACGAGATGAGCCCGCCCTGCAATGCAAGATCTCCGTGAGTACG 837  
671 GCCAAGCTGCCACACGAGATGAGCCCGCCCTGCAATGCAAGATCTCCGTGAGTACG 612  
838 GACGCTTCCACACCGAGGCTTCTGCAATGCGCCGCTGAAGAGATGAGCCCTATG 897  
611 GACGCTTCCACACCGAGGCTTCTGCAATGCGCCGCTGAAGAGATGAGCCCTATG 552  
898 ATGATCTGCGCCCTTGCAGCCCGCCCGCCAGCGTCAACAGTACCTGAAGTACCAAGGCC 957  
551 ATGATCTGCGCCCTTGCAGCCCGCCCGCCAGCGTCAACAGTACCTGAAGTACCAAGGCC 492  
958 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 1017  
491 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 432  
1018 GATGCTGGGGGGGGGCTTGGCCAGTGCCTTCAAGGAGTGGGCGCCCAAGTGCCTGTC 1077  
431 GATGCTGGGGGGGGGCTTGGCCAGTGCCTTCAAGGAGTGGGCGCCCAAGTGCCTGTC 372  
1078 GCATCTCCCAACCGAGGCGCCCGCCAGCGTCTTGTTCACAGACTTAAAGCAACAGCTGG 1137  
371 GCATCTCCCAACCGAGGCGCCCGCCAGCGTCTTGTTCACAGACTTAAAGCAACAGCTGG 312  
1138 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 1197  
311 AAGGCTCTTCCGGAAGGCGCCAGGCGCGTGAAGAGACAAAGTGCACCATCCAGTGAAGCG 252  
1198 CTGTGCCAAGCTCTTGTGTTTCAATCTCTCTGTGGAGAGACAACTCTCTGAGCT 1257  
251 CTGTGCCAAGCTCTTGTGTTTCAATCTCTCTGTGGAGAGACAACTCTCTGAGCT 192  
1258 CAAGATTAGGAGAGACTCAAGTTCACCTCTCTCTGAGGTTGGAGAAATGTTGA 1317  
191 CAAGATTAGGAGAGACTCAAGTTCACCTCTCTCTGAGGTTGGAGAAATGTTGA 132  
1318 TGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACCTCTGAGCAAGCTCAGAGTGA 1377  
131 TGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACCTCTGAGCAAGCTCAGAGTGA 72  
1378 AAAGGACACAGAAAGGCGCAGATGAGAAAGTCTCTCTCTGAGGATTAACCCGAGCTTG 1437  
71 AAAGGACACAGAAAGGCGCAGATGAGAAAGTCTCTCTCTGAGGATTAACCCGAGCTTG 12  
1438 GTTTGGG 1444

Db 11 GTTTGGG 5

RESULT 11  
US-10-027-632-25125/c  
; Sequence 25125, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO: 25125  
; LENGTH: 753  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-25125

Query March 18.6%; Score 574; DB 13; Length 753;

Best Local Similarity 99.6%; Pred. No. 2.3e-265; Matches 724; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 718 GAGGTGTGCGGCAAGAAACACCAACGATGATCTACGTCTCTTCAAGCATG 777  
Db 731 GAGGTGTGCGGCAAGAAACACCAACGATGATCTACGTCTCTTCAAGCATG 672  
Qy 778 GCCAAGCTGCCACAGAGATGAGCCCGCTTCATGCAAGATCTCCGTCAATGCGT 837  
Db 671 GCCAAGCTGCCACAGAGATGAGCCCGCTTCATGCAAGATCTCCGTCAATGCGT 612  
Qy 838 GAGCGCTTCCACCCCAAGGCGCTTTCGATGCGCGCGTCAAGAGATGAGCGCTATGCG 897  
Db 611 GAGCGCTTCCACCCCAAGGCGCTTTCGATGCGCGCGTCAAGAGATGAGCGCTATGCG 552  
Qy 898 ATGGTCTGCGCTTTCGCGCGCGCGCGCGCGCTCAAGATGAGCGCTATGAGGCG 957  
Db 551 ATGGTCTGCGCTTTCGCGCGCGCGCGCGCGCTCAAGATGAGCGCTATGAGGCG 492  
Qy 958 AAGGTCTTCGGAAGGCGCAGGCGCGTGAAGAGGCAAGTGCACCATCAGTGAAGGAG 1017  
Db 491 AAGGTCTTCGGAAGGCGCAGGCGCGTGAAGAGGCAAGTGCACCATCAGTGAAGGAG 432  
Qy 1018 GATGCTGGGGGGGGCTTGGCGATGCTTTCAGGGAAGTGGCCCAATGCGCACTGTC 1077  
Db 431 GATGCTGGGGGGGGCTTGGCGATGCTTTCAGGGAAGTGGCCCAATGCGCACTGTC 372  
Qy 1078 GATGCTGGGGGGGGCTTGGCGATGCTTTCAGGGAAGTGGCCCAATGCGCACTGTC 1137  
Db 371 GATGCTGGGGGGGGCTTGGCGATGCTTTCAGGGAAGTGGCCCAATGCGCACTGTC 312  
Qy 1138 AGGCGCGCGGGCGCTGCGCTTCGCAATTCGCTTCCTTTCAGGCTTTCTTGAAGTCG 1197  
Db 311 AGGCGCGCGGGCGCTGCGCTTCGCAATTCGCTTCCTTTCAGGCTTTCTTGAAGTCG 252

Qy 1198 CTGTGCCACAGCTCTGTGTGTTTCACTCTCTGTGGAGAGACATCTTCAGCCT 1257  
Db 251 CTGTGCCACAGCTCTGTGTGTTTCACTCTCTGTGGAGAGACATCTTCAGCCT 192  
Qy 1258 CAAGATTGAGGCGAGAGCTAAGTTACACCTTCTCTCTGCGGTTGGAAGAAATGTTGA 1317  
Db 191 CAAGATTGAGGCGAGAGCTAAGTTACACCTTCTCTCTGCGGTTGGAAGAAATGTTGA 132  
Qy 1318 TGCCAGAGGGGTGAGATTCGTGCTATATGAGCCCTCTGCGGACAGCCTGAGATGA 1377  
Db 131 TGCCAGAGGGGTGAGATTCGTGCTATATGAGCCCTCTGCGGACAGCCTGAGATGA 72  
Qy 1378 AAGGACACAGAAAGGCGAGATGAGAAAGTCTCTCTCTGCGGATTAACCCAGCTTG 1437  
Db 71 AAGGACACAGAAAGGCGAGATGAGAAAGTCTCTCTCTGCGGATTAACCCAGCTTG 12  
Qy 1438 GTTTGGG 1444  
Db 11 GTTTGGG 5

RESULT 12  
US-10-027-632-25123/c  
; Sequence 25123, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO: 25123  
; LENGTH: 753  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-25123

Query March 18.6%; Score 574; DB 16; Length 753;

Best Local Similarity 99.6%; Pred. No. 2.3e-265; Matches 724; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 718 GAGGTGTGCGGCAAGAAACACCAACGATGATGTTCTACGTCTCTTCAAGCATG 777  
Db 731 GAGGTGTGCGGCAAGAAACACCAACGATGATGTTCTACGTCTCTTCAAGCATG 672  
Qy 778 GCCAAGCTGCCACAGAGATGAGCCCGCTTCATGCAAGATCTCCGTCAATGCGT 837  
Db 671 GCCAAGCTGCCACAGAGATGAGCCCGCTTCATGCAAGATCTCCGTCAATGCGT 612  
Qy 838 GAGCGCTTCCACCCCAAGGCGCTTTCGATGCGCGCGTCAAGAGATGAGCGCTATGCG 897  
Db 611 GAGCGCTTCCACCCCAAGGCGCTTTCGATGCGCGCGTCAAGAGATGAGCGCTATGCG 552  
Qy 898 ATGGTCTGCGCTTTCGCGCGCGCGCGCGCTCAAGATGAGCGCTATGAGGCG 957  
Db 551 ATGGTCTGCGCTTTCGCGCGCGCGCGCGCGCTCAAGATGAGCGCTATGAGGCG 492



PRIOR FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/167,363  
 PRIOR FILING DATE: 1999-11-23  
 PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 NUMBER OF SEQ ID NOS: 325720  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 25125  
 LENGTH: 753  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-25125

Query Match 18.6%; Score 574; DB 16; Length 753;

Best Local Similarity 99.6%; Pred. No. 2,3e-265;  
 Matches 724; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

718 GAGGTGCGCCCAAGAAACACCAACGATGATGTTCTAGTGTCTTTCAGCATG 777  
 731 GAGGTGCGCCCAAGAAACACCAACGATGATGTTCTAGTGTCTTTCAGCATG 672  
 778 GCCAAGCTGCCACAGAGATGAGCCCGCTTCATGCAAGATCTCCGTGAGTACGAT 837  
 671 GCCAAGCTGCCACAGAGATGAGCCCGCTTCATGCAAGATCTCCGTGAGTACGAT 612  
 838 GAGCCTTCCACCCCAAGGCTTCTTCATGCGCGCGCTCAAGAGATGAGCGCTATGCG 897  
 611 GACCCCTTCCACCCCAAGGCTTCTTCATGCGCGCGCTCAAGAGATGAGCGCTATGCG 552  
 898 ATGATCTGCGCTTTCG 957  
 551 ATGATCTGCGCTTTCG 492  
 958 AAGGTCTTCCGGAAGGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017  
 491 AAGGTCTTCCGGAAGGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 432  
 1018 GATCTGCG 1077  
 431 GATCTGCG 372  
 1078 GATCTGCG 1137  
 371 GATCTGCG 312  
 1138 AAGCCTTCCGGAAGGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197  
 311 AAGCCTTCCGGAAGGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 252  
 1198 CTGTGTCACAGCTCTGTTGTTTCACTCTCTGTTGGAAGACACATCTCTGAGCTT 1257  
 251 CTGTGTCACAGCTCTGTTGTTTCACTCTCTGTTGGAAGACACATCTCTGAGCTT 192  
 1258 CAAGATTAGGCAAGATCAAGTTACCTTCTCTCTGTTGGAAGAAATGTTGA 1317  
 191 CAAGATTAGGCAAGATCAAGTTACCTTCTCTCTGTTGGAAGAAATGTTGA 132  
 1318 TGCCAGAGGGGTGAGATGCTGCTCATATGAGGCTCTGAGGACAGGCTCAGGATGA 1377  
 131 TGCCAGAGGGGTGAGATGCTGCTCATATGAGGCTCTGAGGACAGGCTCAGGATGA 72  
 1378 AAAGGACACAGAGGCGCAGATGAGAAAGGTCTCTCTCTGAGTAAACCCAGCTTG 1437  
 71 AAAGGACACAGAGGCGCAGATGAGAAAGGTCTCTCTCTGAGTAAACCCAGCTTG 12  
 1438 GTTTGGG 1444  
 11 GTTTGGG 5

RESULT 15

US-10-027-632-106883  
 Sequence 106883, Application US/10027632  
 Publication No. US2002019837A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, David G.  
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 TITLE OF INVENTION: Polymorphisms in the Human Genome  
 FILE REFERENCE: 108827,129  
 CURRENT APPLICATION NUMBER: US/10/027,632  
 CURRENT FILING DATE: 2002-04-30  
 PRIOR APPLICATION NUMBER: US 60/219,006  
 PRIOR FILING DATE: 2000-07-12  
 PRIOR APPLICATION NUMBER: US 60/198,676  
 PRIOR FILING DATE: 2000-04-20  
 PRIOR APPLICATION NUMBER: US 60/193,483  
 PRIOR FILING DATE: 2000-03-29  
 PRIOR APPLICATION NUMBER: US 60/185,218  
 PRIOR FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/167,363  
 PRIOR FILING DATE: 1999-11-23  
 PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 NUMBER OF SEQ ID NOS: 325720  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 106883  
 LENGTH: 604  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-106883

Query Match 17.9%; Score 553; DB 13; Length 604;

Best Local Similarity 99.8%; Pred. No. 3e-255;  
 Matches 603; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

846 CCACCCAGGCGCTTTCGATGAGCGCGGTCAAGAGATGAGCGCTATGAGATGATGTC 905  
 1 CCACCCAGGCGCTTTCGATGAGCGCGGTCAAGAGATGAGCGCTATGAGATGATGTC 60  
 906 GCCCTTGGCCCGCCCGCCCAAGCTCAACAGTACCTCAAGTATCAAGGCGCAAGTCT 965  
 61 GCCCTTGGCCCGCCCGCCCAAGCTCAACAGTACCTCAAGTATCAAGGCGCAAGTCT 120  
 966 TCGGAAAGGCGCGCGCGGTGAGAGGACAGTGCACATTCAGTGAAGGAGATGCTGG 1025  
 121 TCGGAAAGGCGCGCGCGGTGAGAGGACAGTGCACATTCAGTGAAGGAGATGCTGG 180  
 1026 GCGGCGGCTTGGCCAGGCGCTTTCAGGAGGAGTGGCCCAAGATGCCCATGTCGATGCC 1085  
 181 GCGGCGGCTTGGCCAGGCGCTTTCAGGAGGAGTGGCCCAAGATGCCCATGTCGATGCC 240  
 1086 CCACCGAGGCGCGCGCGAGAGTCTTGTTCACAGACTTATGAGCACAGACTGAGGCGCCC 1145  
 241 CCACCGAGGCGCGCGCGAGAGTCTTGTTCACAGACTTATGAGCACAGACTGAGGCGCCC 300  
 1146 GCGGCGTGGCTCCGCAATTCGTCGCTTTCACAGACTTCTGTAAGTCCGCTTGTCCA 1205  
 301 GCGGCGTGGCTCCGCAATTCGTCGCTTTCACAGACTTCTGTAAGTCCGCTTGTCCA 360  
 1206 CAGTCTTGGTGGTTTCACTCTCTGTTGGAAGACACATCTGAGGCTCAGAAATTT 1265  
 361 CAGTCTTGGTGGTTTCACTCTCTGTTGGAAGACACATCTGAGGCTCAGAAATTT 420  
 1266 AGCAGAGACTCAAGTTACACTTCTCTCTGAGGTTGGAAGAAATGTTAGTCCAGAG 1325  
 421 AGCAGAGACTCAAGTTACACTTCTCTCTGAGGTTGGAAGAAATGTTAGTCCAGAG 480  
 1326 GGTGAGGATGCTGCGCTCATATGAGGCTTCTGGAACAAGGCTCAGAGTAAAGGACA 1385  
 481 GGTGAGGATGCTGCGCTCATATGAGGCTTCTGGAACAAGGCTCAGAGTAAAGGACA 540  
 1386 CAGAGGCGCAGATGAGAAAGGTCTCTCTCTGAGTAAACCCAGCTTGGT 1445

**Tue Jun 22 10:07:19 2004**

us-09-778-963a-1.oli.rnpb

Page 16

```

Db      541 CAGAAAGCCAGATGAGAAAGTCTCTCTCTGCAATPACACCAAGCTTGTTGGGT 600
Oy      1446 GGCA 1449
Db      601 GGCA 604

```

Search completed: June 20, 2004, 20:45:32  
Job time : 1260 secs





QY	2664	AAGGTCCTCCTAAGTATTGTGTACACAGTCCCATGTATATTAACAAGTACACAGATGT	2723
Db	240	GAGGTCCTCCTGTAGTATTGTGTATGATGCACATGCGCATGTATATTAACAAGTACACAGATGT	299
QY	2724	AACACAGATGTACATCATGTCTCCAGCCCGAGCTGTGCATTAACCTGCACCTGCACCCGACC	2783
Db	300	AACACAGATGTACACATGTCTCCAGCCCGAGCTGTGCATTAACCTGCACCTGCACCCGACC	359
QY	2784	TTGGCCCTGACTG	2797
Db	360	TTGGCCCTGACTG	373

```

RESULT 2
US-09-601-537-10
/ Sequence 10, Application us/09601537
/ Patent No. 6582691
/ GENERAL INFORMATION:
/ APPLICANT: Gallett, Karl-Christian
/ APPLICANT: Mullner, Stefan
/ APPLICANT: Huls, Christoph
/ APPLICANT: Bohmisch, Britta
/ TITLE OF INVENTION: Expression Vector for the Production of
/ TITLE OF INVENTION: Dead Proteins
/ FILE REFERENCE: 50186/002001
/ CURRENT APPLICATION NUMBER: US/09/601,537
/ CURRENT FILING DATE: 2000-08-02
/ PRIOR APPLICATION NUMBER: PCT/EP99/00829
/ PRIOR FILING DATE: 1999-02-09
/ PRIOR APPLICATION NUMBER: DE 198 05 781.4
/ PRIOR FILING DATE: 1998-02-12
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10
/ LENGTH: 441
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: p153-NT3
US-09-601-537-10

```

[illegible]

RESULT 3  
US-09-461-325-44  
Sequence 44, Application US/09461325A  
Patent No. 6475753  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 94 Human Secreted Proteins  
FILE REFERENCE: P2029P1  
CURRENT APPLICATION NUMBER: US/09/461,325A  
CURRENT FILING DATE: 1999-12-14  
EARLIER APPLICATION NUMBER: PCT/US99/13418  
EARLIER FILING DATE: 1999-06-15  
EARLIER APPLICATION NUMBER: 60/089,507  
EARLIER FILING DATE: 1998-06-16  
EARLIER APPLICATION NUMBER: 60/089,508  
EARLIER FILING DATE: 1998-06-16  
EARLIER APPLICATION NUMBER: 60/089,509  
EARLIER FILING DATE: 1998-06-16

```

? EARLIER APPLICATION NUMBER: 60/089,510
?
? EARLIER FILING DATE: 1998-06-16
?
? EARLIER APPLICATION NUMBER: 60/090,112
?
? EARLIER FILING DATE: 1998-06-22
?
? EARLIER APPLICATION NUMBER: 60/090,113
?
? EARLIER FILING DATE: 1998-06-22
?
? NUMBER OF SEQ ID NOS: 532
?
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 44
?
? LENGTH: 569
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
US-09-461-325-44

```

Query Match 3.5%; Score 108; DB 4; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.8e-31;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	22975	30382
Db	433	492
QY	3035	3082
Db	493	540

```

RESULT 4
US-10-012-542-44
Sequence 44, Application US/10012542
Patent No. 6627741
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 94 Human Secreted Proteins
FILE REFERENCE: P202921
CURRENT APPLICATION NUMBER: US/10/012,542
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
NUMBER OF SEQ ID NOS: 532
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 44
LENGTH: 569
TYPE: DNA
ORGANISM: Homo sapiens
IS-10-012-542--44

```

[illegible]

## RESULT 5

```

US-09-620-405B-465
: Sequence 465, Application US/09620405B
: Patent No. 6528054
:
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
:
: TITLE OF INVENTION: COMPOSITIONS AND METHODS AND
: TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
:
: FILE REFERENCE: 210121.470C8
:
: CURRENT APPLICATION NUMBER: US/09/620,405B
:
: CURRENT FILING DATE: 2000-07-20
:
: NUMBER OF SEQ ID NOS: 495
:
: SOFTWARE: FastSeq for Windows Version 3.0
:
: SEQ ID NO 465
:
: LENGTH: 674
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: US-09-620-405B-465

```

	Query Match	3.5%;	Score 108;	DB 4;	Length 674;
	Best Local Similarity	100.0%;	Pred. No. 1.7e-31;		
	Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2975	CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	303
Dd	549	CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	608
QY	3035	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3082
Dd	609	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	656

RESULT 6  
 US-09-433-826B-465  
 ; Sequence 465, Application US/09433826B  
 ; Patent No. 6579873  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Harlocker, Susan L.  
 ; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.470C4  
 ; CURRENT APPLICATION NUMBER: US/09/433.826B  
 ; CURRENT FILING DATE: 1999-11-03  
 ; NUMBER OF SEQ ID NOS: 474  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 465  
 ; LENGTH: 674  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-433-826B-465

[illegible]

## RESULT 7

```

US-09-604-287A-465
: Sequence 465, Application US/09604287A
: Patent No. 6586572
:
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugui
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiaqichun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
: FILE REFERENCE: 210121.470C7
: CURRENT APPLICATION NUMBER: US/09/604,287A
: CURRENT FILING DATE: 2000-06-22
: NUMBER OF SEQ ID NOS: 489
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 465
: LENGTH: 674
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-604-287A-465

```

[illegible]

```

RESULT 8
US-09-834-759-465
: Sequence 465: Application US/09834759
: Patent No. 6680197
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yuguu
: APPLICANT: Dillon, Devin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.470C9
: CURRENT APPLICATION NUMBER: US/09/834,759
: CURRENT FILING DATE: 2001-04-13
: NUMBER OF SEQ ID NOS: 547
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 465
: LENGTH: 674
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-834-759-465

```

[illegible]

RESULT 9  
US-09-073-569-1  
; Sequence 1, Application US/09073569  
; Patent No. 6084088  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; TITLE OF INVENTION: NOVEL TUMOR ANTIGENS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 1201 Eastlake Avenue East  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/073,569  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sawislak, Deborah A  
; REGISTRATION NUMBER: 37,438  
; REFERENCE/DOCKET NUMBER: 97-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-442-6672  
; TELEFAX: 206-442-6678  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1733 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 34...1344  
; OTHER INFORMATION:  
; US-09-073-569-1  
Query Match 3.5%; Score 108; DB 3; Length 1733;  
Best Local Similarity 100.0%; Pred. No. 1.5e-31;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2975 CAAA 3034  
DB 1607 CAAA 1666  
QY 3035 AA 3082  
DB 1667 AA 1714  
RESULT 10  
US-09-071-224-3  
; Sequence 3, Application US/09071224  
; Patent No. 6271343  
; GENERAL INFORMATION:  
; APPLICANT: Lok, Si  
; APPLICANT: Presnell, Scott R.  
; APPLICANT: Jelmberg, Anna C.  
; APPLICANT: Gilbert, Teresa  
; APPLICANT: Foster, Donald C.  
; APPLICANT: Adams, Robyn L.

APPLICANT: Lehner, Joyce M.  
; TITLE OF INVENTION: MAMMALIAN ZCYTORS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: ZymoGenetics  
; STREET: 1201 Eastlake Ave East  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,224  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lum, Paul G  
; REGISTRATION NUMBER: 32,743  
; REFERENCE/DOCKET NUMBER: 96-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-442-6627  
; TELEFAX: 206-442-6678  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1813 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 88...1362  
; OTHER INFORMATION:  
; US-09-071-224-3  
Query Match 3.5%; Score 108; DB 3; Length 1813;  
Best Local Similarity 100.0%; Pred. No. 1.5e-31;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2975 CAAA 3034  
DB 1687 CAAA 1746  
QY 3035 AA 3082  
DB 1747 AA 1794  
RESULT 11  
US-09-907-794A-212  
; Sequence 212, Application US/09907794A  
; Patent No. 6635468  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Balon, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Billen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.

```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavin, Ivar J.
: APPLICANT: Macher, Jennie P.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William, I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: 10466-14
: CURRENT APPLICATION NUMBER: US/09/907,794A
: PRIOR FILING DATE: 2001-07-17
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: US 60/143,048
: PRIOR FILING DATE: 1999-07-07
: PRIOR APPLICATION NUMBER: US 60/145,698
: PRIOR FILING DATE: 1999-07-26
: PRIOR APPLICATION NUMBER: US 60/146,222
: PRIOR FILING DATE: 1999-07-28
: PRIOR APPLICATION NUMBER: PCT/US99/20594
: PRIOR FILING DATE: 1999-09-08
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIOR FILING DATE: 1999-09-13
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/21547
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/23089
: PRIOR FILING DATE: 1999-10-05
: PRIOR APPLICATION NUMBER: PCT/US99/28214
: PRIOR FILING DATE: 1999-11-29
: PRIOR APPLICATION NUMBER: PCT/US99/28313
: PRIOR FILING DATE: 1999-11-30
: PRIOR APPLICATION NUMBER: PCT/US99/28564
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/28565
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/30095
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: PCT/US99/30911
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US99/30999
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US00/00219
: PRIOR FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 423
: SEQ ID NO 212
: LENGTH: 1985
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-907-794A-212

Query Match      3.5%; Score 108; DB 4; Length 1985;
Best Local Similarity 100.0%; Pred. No. 1.5e-31;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2975 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3034
Db 1958 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1917
Qy 3035 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3082
Db 1918 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1965

RESULT 12
```

```

US-09-905-125A-212
: Sequence 212, Application US/09905125A
: Patent No. 6664376
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertschen, Mary E.
: APPLICANT: Goddard, A.
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavin, Ivar J.
: APPLICANT: Macher, Jennie P.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William, I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: 10466-14
: CURRENT APPLICATION NUMBER: US/09/905,125A
: PRIOR FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: US 60/143,048
: PRIOR FILING DATE: 1999-07-07
: PRIOR APPLICATION NUMBER: US 60/145,698
: PRIOR FILING DATE: 1999-07-26
: PRIOR APPLICATION NUMBER: US 60/146,222
: PRIOR FILING DATE: 1999-07-28
: PRIOR APPLICATION NUMBER: PCT/US99/20594
: PRIOR FILING DATE: 1999-09-08
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIOR FILING DATE: 1999-09-13
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/21547
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/23089
: PRIOR FILING DATE: 1999-10-05
: PRIOR APPLICATION NUMBER: PCT/US99/28214
: PRIOR FILING DATE: 1999-11-29
: PRIOR APPLICATION NUMBER: PCT/US99/28313
: PRIOR FILING DATE: 1999-11-30
: PRIOR APPLICATION NUMBER: PCT/US99/28564
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/28565
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/30095
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: PCT/US99/30911
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US99/30999
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US00/00219
: PRIOR FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 423
: SEQ ID NO 212
: LENGTH: 1985
: TYPE: DNA
: ORGANISM: Homo sapiens
```

US-09-905-125A-212

Query Match 3.5%; Score 108; DB 4; Length 1985;  
Best Local Similarity 100.0%; Pred. No. 1.5e-31;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2975 CAAA 3034  
DB 1858 CAAA 1917  
QY 3035 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3082  
DB 1918 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1965

RESULT 13  
US-09-902-775A-212

Sequence 212, Application US/09902775A

Patent No. 6686451

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerltzen, Mary E.

APPLICANT: Goddard, A.

APPLICANT: Grimaldi, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth, J.

APPLICANT: Kijavlin, Ivar J.

APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/902,775A

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: US 60/146,222

PRIOR FILING DATE: 1999-07-28

PRIOR APPLICATION NUMBER: PCT/US99/20594

PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-13

PRIOR APPLICATION NUMBER: PCT/US99/21090

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/23089

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: PCT/US99/28214

PRIOR FILING DATE: 1999-11-29

PRIOR APPLICATION NUMBER: PCT/US99/28313

PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: PCT/US99/28564

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/28565

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/30095

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: PCT/US99/30911

PRIOR FILING DATE: 1999-12-20

PRIOR APPLICATION NUMBER: PCT/US99/30999

PRIOR FILING DATE: 1999-12-20

PRIOR APPLICATION NUMBER: PCT/US00/00219

PRIOR FILING DATE: 2000-01-05

NUMBER OF SEQ ID NOS: 423

SEQ ID NO 212

LENGTH: 1985

TYPE: DNA

ORGANISM: Homo sapiens

US-09-902-775A-212

Query Match 3.5%; Score 108; DB 4; Length 1985;  
Best Local Similarity 100.0%; Pred. No. 1.5e-31;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2975 CAAA 3034  
DB 1858 CAAA 1917  
QY 3035 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3082  
DB 1918 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1965

RESULT 14  
US-08-785-310A-2

Sequence 2, Application US/08785310A

Patent No. 5840532

GENERAL INFORMATION:

APPLICANT: McKnight, Steven L.

APPLICANT: Russell, David W.

TITLE OF INVENTION: Neuronal PAS Domain Protein

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESS: SCIENCE &amp; TECHNOLOGY LAW GROUP

STREET: 268 BUSH STREET, SUITE 3200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/785,310A

FILING DATE: 21-JAN-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: UTSD:1226

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2082 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-785-310A-2

Query Match 3.5%; Score 108; DB 2; Length 2082;



THIS PAGE BLANK (USPTO)



RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
B0954076	947 bp	mRNA	linear	EST	21-AUG-2002							
AGNCOURT_8866147	NIH_MGC_101	Homo sapiens cDNA clone IMAGE:646413	5', mRNA sequence.	B0954076	B0954076.1	GI:22369554	EST.					
							Homo sapiens (human)					
							Homo sapiens					
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
							Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
							1 (bases 1 to 947)					
							NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .					
							National Institutes of Health, Mammalian Gene Collection (MGC)					
							Unpublished (1999)					
							Contact: Robert Strausberg, Ph.D.					
							Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>					
							Tissue Procurement: ATCC					
							cDNA Library Preparation: Rubin Laboratory					
							cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)					
							DNA Sequencing by: Agencourt Bioscience Corporation					
							Clone distribution: MGC clone distribution information can be					
							found through the I.M.A.G.E. Consortium/LNL at:					
							<a href="http://image.lnl.gov">http://image.lnl.gov</a>					
							Plate: LUCM2656 row: 1 column: 24					
							High quality sequence stop: 668.					

FEATURES  
source

Location/Qualifiers  
1. 947  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6464135"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NH\_MGC\_101"  
/note="Organ: lung; Vector: pOT7; Site\_1: EcorI; Site\_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcorI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NH\_MGC Library."

## ORIGIN

Query Match 23.4%; Score 722; DB 13; Length 947;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 772; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
183 CCGGCGCCGCAATCCAGCCCGGAGCCATGATGAAGACTTTGTCAGCGGAGACTGCAC 242
1 CCGGCGCCGCAATCCAGCCCGGAGCCATGATGAAGACTTTGTCAGCGGAGACTGCAC 60
243 GCTCAGTGGCCCGGCAAAATCATACCGGATGAGTGTGCTGGGTGCTCTCGGGTGG 302
61 GCTCAGTGGCCCGGCAAAATCATACCGGATGAGTGTGCTGGGTGCTCTCGGGTGG 120
303 CAAGAGCTCATCTGTCGTCTGCTCTCATGAGCGCTTTGAGAGCAGATACACACCCAC 362
121 CAAGAGCTCATCTGTCGTCTGCTCTCATGAGCGCTTTGAGAGCAGATACACACCCAC 180
363 CATGAGAGCTTCCACCGTAAAGTATACAAATCCGCGGCGCATGTACAGCTCGACAT 422
181 CATGAGAGCTTCCACCGTAAAGTATACAAATCCGCGGCGCATGTACAGCTCGACAT 240
423 CCTGATCTCTGTCGACACACCCCTCCCGGCGCATGAGTGTGCTGGGTGCTCTCGGGTGG 482
241 CCTGATCTCTGTCGACACACCCCTCCCGGCGCATGAGTGTGCTGGGTGCTCTCGGGTGG 300
483 GGATGTCCTCATCTGTCGTCTGCTCTCATGAGCGCTTTGAGAGCAGATACACACCCAC 542
301 GGATGTCCTCATCTGTCGTCTGCTCTCATGAGCGCTTTGAGAGCAGATACACACCCAC 360
543 CCTTCAAGAGCAGATCTGAGAGTCAAGTCTGCTGTAAGAACAGACCAAGAGAGCGGC 602
361 CCTTCAAGAGCAGATCTGAGAGTCAAGTCTGCTGTAAGAACAGACCAAGAGAGCGGC 420
603 GGAGTGGCCCATGTCATCTGTGGCAACAAGAACCAACGAGGAGTGTGCGGCAAGT 662
421 GGAGTGGCCCATGTCATCTGTGGCAACAAGAACCAACGAGGAGTGTGCGGCAAGT 480
663 GCCCAACCAACGAGGCGAGCTGTGTGTGGGCGACAGAGAACTCCGCTTACAGAGT 722
481 GCCCAACCAACGAGGCGAGCTGTGTGTGGGCGACAGAGAACTCCGCTTACAGAGT 540
723 GTGGCCCAAGAAACACCAAGTGAAGAGATTTTCAAGTGTCTTCAAGATGGCCAA 782
DB GTGGCCCAAGAAACACCAAGTGAAGAGATTTTCAAGTGTCTTCAAGATGGCCAA 600
541 GTGGCCCAAGAAACACCAAGTGAAGAGATTTTCAAGTGTCTTCAAGATGGCCAA 600
783 GCTGCAACAGAGATGAGCGCCGCTGATCGAAGATCTCGTGAAGTACAGTGAAGC 842
DB GCTGCAACAGAGATGAGCGCCGCTGATCGAAGATCTCGTGAAGTACAGTGAAGC 660
601 GCTGCAACAGAGATGAGCGCCGCTGATCGAAGATCTCGTGAAGTACAGTGAAGC 660
843 CTTCCACCCCAAGGCTTCTGATGCGCGCTCAAGAGATGAGAGCGCTTACAGTGT 902
DB CTTCCACCCCAAGGCTTCTGATGCGCGCTCAAGAGATGAGAGCGCTTACAGTGT 720
903 CTTCCACCCCAAGGCTTCTGATGCGCGCTCAAGAGATGAGAGCGCTTACAGTGT 955
```

DB 721 CTCGCCCTTCCGCCGCCGCCGCCAGGCTCAACAGTACTCAAGTACATCAAG 773

RESULT 2  
BX394422/c 974 bp mRNA linear EST 13-MAY-2003  
LOCUS BX394422 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
DEFINITION CDNA clone CS0DC015YK10.3-PRIME, mRNA sequence.  
ACCESSION BX394422.1 GI:30631374  
VERSION BX394422.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 974)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
10498.f for more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DC015BF05NP1&cluster=10498.f. Contact :  
Peng Liang Email : liliang@life.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DC015BF05NP1.

FEATURES  
source

Location/Qualifiers  
1. 974  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC015YK10"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and Ecor V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 23.2%; Score 716; DB 13; Length 974;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 766; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
2001 GGGACACATGATGAGACAGGCGCCCACTTTGCTCCCTAGCAAAATTCAAGGTACAGCTCA 2060
962 GGGACACATGATGAGACAGGCGCCCACTTTGCTCCCTAGCAAAATTCAAGGTACAGCTCA 903
2061 CTTAGAACAGGCTGCTCTTACTGTGTGCTCTTCTTCAAGCATTTTAAAGACCTACTG 2120
902 CTTAGAACAGGCTGCTCTTACTGTGTGCTCTTCTTCAAGCATTTTAAAGACCTACTG 843
2121 GGTGTGGGTTCATCTGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2180
DB GGTGTGGGTTCATCTGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 783
842 GGTGTGGGTTCATCTGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 783
2181 CCGCTGCTGCCCAACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2240
DB CCGCTGCTGCCCAACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 723
782 CCGCTGCTGCCCAACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 723
2241 GGACCCAGAACTGAGCTTGGAGGAGATCCAGAGAAAAGCTCAAGGGGCGGTCTTCTCTT 2300
DB GGACCCAGAACTGAGCTTGGAGGAGATCCAGAGAAAAGCTCAAGGGGCGGTCTTCTCTT 663
722 GGACCCAGAACTGAGCTTGGAGGAGATCCAGAGAAAAGCTCAAGGGGCGGTCTTCTCTT 663
2301 GTGGCCGGGATTTGGGCTATGCTGGGTACCAACATGTAAGTCAAGGATGAGTGGTTTGAAC 2360
DB GTGGCCGGGATTTGGGCTATGCTGGGTACCAACATGTAAGTCAAGGATGAGTGGTTTGAAC 603
```

QY 2361 CCATTAACCAAGGCGCTTGTATCATGCTCTTAACAATATTTTGTATTTAATCTCT 2420  
 |||||  
 Db 602 CCAATTAACCAAGGCGCTTGTATCATGCTCTTAACAATATTTTGTATTTAATCTCT 543  
 |||||  
 QY 2421 CTAACATATTTGAATTTTAAAGGCGCTTAAGAACTTAAGTATCTTCTATTTGGGCTTTTC 2480  
 |||||  
 Db 542 CTAACATATTTGAATTTTAAAGGCGCTTAAGAACTTAAGTATCTTCTATTTGGGCTTTTC 483  
 |||||  
 QY 2481 TGAAGTTCAAGAGGAGTAAATGATCTCTCCAGGTCAACAGCAAGTCTGTGGGTGGAG 2540  
 |||||  
 Db 482 TGAAGTTCAAGAGGAGTAAATGATCTCTCCAGGTCAACAGCAAGTCTGTGGGTGGAG 423  
 |||||  
 QY 2541 AAGCAAGTACGCTGGGCACTTCAATATACCAAGATGTGCTCCCTCTTGAATGCTTG 2600  
 |||||  
 Db 422 AAGCAAGTACGCTGGGCACTTCAATATACCAAGATGTGCTCCCTCTTGAATGCTTG 363  
 |||||  
 QY 2601 GCCCTGGGCGCTTCAAGGCTTTGGAGACATCTTGTCTCAACCTCTCCCTAGATCACTC 2660  
 |||||  
 Db 362 GCCCTGGGCGCTTCAAGGCTTTGGAGACATCTTGTCTCAACCTCTCCCTAGATCACTC 303  
 |||||  
 QY 2661 TGTAGAGGCTCCCTGTATATTTGTGTATACCAATGCCCATGTATATATACACAGCA 2720  
 |||||  
 Db 302 TGTAGAGGCTCCCTGTATATTTGTGTATACCAATGCCCATGTATATATACACAGCA 243  
 |||||  
 QY 2721 TGTACACACAGATGTATACATGCTCCAGGCCAGCTGTGATACCTG 2767  
 |||||  
 Db 242 TGTACACACAGATGTATACATGCTCCAGGCCAGCTGTGATACCTG 196  
 |||||

RESULT 3  
 LOCUS B0073742 1006 bp mRNA linear EST 02-APR-2002  
 DEFINITION AGENCOURT 7046611 NIH\_MGC\_101 Homo sapiens CDNA clone IMAGE:5806628  
 5', mRNA sequence.  
 B0073742  
 B0073742.1 GI:19902788  
 EST.  
 ACCESSION  
 VERSION B0073742.1  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 1006)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ULNL at:  
 http://image.llnl.gov  
 Plate: LICM2047 row: 1 column: 21  
 High quality sequence stop: 671.  
 Location/Qualifiers  
 1..1006  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5806628"  
 /issue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_101"  
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
 XhoI; CDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

FEATURES  
 source

ORIGIN  
 Query Match 22.9%; Score 707; DB 13; Length 1006;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 757; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 183 CCGGCGCGGCGATTCGCCAGCGCCGAGCATATGAAGACTTTGTCCAGCGGAACTGAC 242  
 |||||  
 Db 1 CCGGCGCGGCGATTCGCCAGCGCCGAGCATATGAAGACTTTGTCCAGCGGAACTGAC 60  
 |||||  
 QY 243 GCTCAGTGTGCGCGCCAACTATACCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 302  
 |||||  
 Db 61 GCTCAGTGTGCGCGCCAACTATACCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120  
 |||||  
 QY 303 CAAGAGCTCCATGT 362  
 |||||  
 Db 121 CAAGAGCTCCATGT 180  
 |||||  
 QY 363 CATGAGGACTTCCACCGTAAAGTATACCAATCGCGCGGAGCATGTACAGCTGCAT 422  
 |||||  
 Db 181 CATGAGGACTTCCACCGTAAAGTATACCAATCGCGCGGAGCATGTACAGCTGCAT 240  
 |||||  
 QY 423 CCTGGAATACCTGTGCAACCAACCCCTTCCCGCATGCGCAGGCTGTCCATCTTCACAG 482  
 |||||  
 Db 241 CCTGGAATACCTGTGCAACCAACCCCTTCCCGCATGCGCAGGCTGTCCATCTTCACAG 300  
 |||||  
 QY 483 GGATGTCTTCACTCTGT 542  
 |||||  
 Db 301 GGATGTCTTCACTCTGT 360  
 |||||  
 QY 543 CCTTCAGAGAGAGATCTCGAGGTCAGAGTCTGCTGCTGAAGAAACAAGCAAGAGAGCGGC 602  
 |||||  
 Db 361 CCTTCAGAGAGAGATCTCGAGGTCAGAGTCTGCTGCTGAAGAAACAAGCAAGAGAGCGGC 420  
 |||||  
 QY 603 GAGAGTGCCTCATGTGTATCTGTGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662  
 |||||  
 Db 421 GAGAGTGCCTCATGTGTATCTGTGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 |||||  
 QY 663 GCCCAACCAAG 722  
 |||||  
 Db 481 GCCCAACCAAG 540  
 |||||  
 QY 723 GTCCGCGCAAG 782  
 |||||  
 Db 541 GTCCGCGCAAG 600  
 |||||  
 QY 783 GCTGCAACAGAGATGAG 842  
 |||||  
 Db 601 GCTGCAACAGAGATGAG 660  
 |||||  
 QY 843 CTTTCAACCCAG 902  
 |||||  
 Db 661 CTTTCAACCCAG 720  
 |||||  
 QY 903 CTTGCGCTTGTGCG 940  
 |||||  
 Db 721 CTTGCGCTTGTGCG 758  
 |||||

RESULT 4  
 LOCUS B1754083 888 bp mRNA linear EST 25-SEP-2001  
 DEFINITION 603027639F1 NIH\_MGC\_114 Homo sapiens CDNA clone IMAGE:5198024 5',  
 mRNA sequence.  
 B1754083  
 B1754083.1 GI:15745661  
 EST.  
 ACCESSION  
 VERSION B1754083  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 888)  
 NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM11495 row: n column: 09  
High quality sequence stop: 796.  
Location/Qualifiers  
1. 888

FEATURES  
source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5198024"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_114"  
/note="Organ: Brain; Vector: pCMV-SPORT6, Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 22.4%; Score 690; DB 12; Length 888;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 740; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

122 CAGCAGGAGCTCCCGCAGCTCCGCGCTTCCAGGAGCTCTGAGCCCTGCAGG 181  
1 CAGCAGGAGCTCCCGCAGCTCCGCGCTTCCAGGAGCTCTGAGCCCTGCAGG 60  
182 GCCCGCGCCGCTTCCAGGAGCTCCGCGCTTCCAGGAGCTCTGAGCCCTGCAGG 241  
61 GCCCGCGCCGCTTCCAGGAGCTCCGCGCTTCCAGGAGCTCTGAGCCCTGCAGG 120  
242 CGCTCAGTGTGCGCCGCAAAAATCATACCGCATGCTGTGCTGCTGCGGTGG 301  
121 CGCTCAGTGTGCGCCGCAAAAATCATACCGCATGCTGTGCTGCTGCGGTGG 180  
302 GCAAGAGCTCATGCTGTCTGCTTCTCAATGCGCGCTTGAAGAGCAGTACACCCA 361  
181 GCAAGAGCTCATGCTGTCTGCTTCTCAATGCGCGCTTGAAGAGCAGTACACCCA 240  
362 CCATCAGAGACTTCCAGCGTAAAGTATACATCCGCGCGCATGTACACAGCTGACA 421  
241 CCATCAGAGACTTCCAGCGTAAAGTATACATCCGCGCGCATGTACACAGCTGACA 300  
422 TCTTGATACCTCTGAGCAACACCCCTTCCCGCATGCGAGGCTGTCCATCTCCAG 481  
301 TCTTGATACCTCTGAGCAACACCCCTTCCCGCATGCGAGGCTGTCCATCTCCAG 360  
482 GGGATGCTTCAATCTGTGTTGAGCTGATTAACCGGAGAGCTCTTCAGATGAGTCAAGC 541  
361 GGGATGCTTCAATCTGTGTTGAGCTGATTAACCGGAGAGCTCTTCAGATGAGTCAAGC 420  
542 GCTTCAGAGAGAGTCTGAGAGTCAAGTCTGCTTGAAGAAACAAAGCAAGAGAGGGG 601  
421 GCTTCAGAGAGAGTCTGAGAGTCAAGTCTGCTTGAAGAAACAAAGCAAGAGAGGGG 480  
602 CGAGCTGCGCATGCTGTGAGCAACAAAGCAAGCAAGAGTGTGCGCGCAGG 661  
481 CGAGCTGCGCATGCTGTGAGCAACAAAGCAAGCAAGAGTGTGCGCGCAGG 540  
662 TGCCACCAACCAAGAGCGAGCTGTGCTGTGCGGCGAGCAAGAACTCCGCTTACAGG 721

Db  
541 TGCCACCAACCAAGAGCGAGCTGTGCTGTGCGGCGAGCAAGTGCCTTACCTGAGG 600  
Qy  
722 TGTCGCGCAAGAGAGAACCAACAGTGAAGAGATTTTACGTGCTTTCAGCATGAGCA 781  
Db  
601 TGTCGCGCAAGAGAGAACCAACAGTGAAGAGATTTTACGTGCTTTCAGCATGAGCA 660  
Qy  
782 AGCTGCCACAGAGATGAGCCCGCTGCAATGCAAGATTTCCGTGAGTACGTTGACG 841  
Db  
661 AGCTGCCACAGAGATGAGCCCGCTGCAATGCAAGATTTCCGTGAGTACGTTGACG 720  
Qy  
842 CTTTCACCGCCGAGGCGCTTCT 862  
Db  
721 CTTTCACCGCCGAGGCGCTTCT 741

RESULT 5  
BQ947936 904 bp mRNA linear EST 21-AUG-2002  
LOCUS BQ947936  
DEFINITION AGERCOURT\_8813043 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6426549  
5', mRNA sequence.  
ACCESSION BQ947936  
VERSION BQ947936.1 GI:22363414  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 904)  
NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contract: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM2610 row: b column: 22  
High quality sequence stop: 669.  
Location/Qualifiers  
1. 904  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6426549"  
/tissue="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_101"  
/note="Organ: lung; Vector: pOT7; Site\_1: EcoRI; Site\_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA Synthesis Kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

FEATURES  
source

## ORIGIN

Query Match 20.8%; Score 641; DB 13; Length 904;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  
183 CCCGCGCCGCAATCCAGCCGCGAGCATGATGAAGACTTTGTCAGCGGAACTGCAC 242  
Db  
1 CCCGCGCCGCAATCCAGCCGCGAGCATGATGAAGACTTTGTCAGCGGAACTGCAC 60  
Qy  
243 GCTCAGTGTGCGCCGCAAAAATCATACCGCATGCTGTGCTGCGGTGG 302  
Db  
61 GCTCAGTGTGCGCCGCAAAAATCATACCGCATGCTGTGCTGCGGTGG 120

QY 303 CAAGAGCTCATGCTGCTGCTTCTCAATGCGCGCTTTGAGACAGTACACACCCAC 362  
DB 121 CAAGAGCTCATGCTGCTGCTTCTCAATGCGCGCTTTGAGACAGTACACACCCAC 180  
QY 363 CATGAGAGCTTCCACCGTAAAGTATACATCCGCGCGGACATGTACAGTCCGACAT 422  
DB 181 CATGAGAGCTTCCACCGTAAAGTATACATCCGCGCGGACATGTACAGTCCGACAT 240  
QY 423 CTTGAGATCTCTGCGAACACACCCCTTCCCGGACATGCGAGGCTGTCTACACAG 482  
DB 241 CTTGAGATCTCTGCGAACACACCCCTTCCCGGACATGCGAGGCTGTCTACACAG 300  
QY 483 GGATGTCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542  
DB 301 GGATGTCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 543 CTTTCAAGAGAGATCTCTGAGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 602  
DB 361 CTTTCAAGAGAGATCTCTGAGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 603 GGAGCTGCTCATGCTCATCTGTGCGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 662  
DB 421 GGAGCTGCTCATGCTCATCTGTGCGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
QY 663 GCGGACACCGAGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722  
DB 481 GCGGACACCGAGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 723 GTCGCGCAAG 782  
DB 541 GTCGCGCAAG 600  
QY 783 GCTGCGCAAG 842  
DB 601 GCTGCGCAAG 660  
QY 843 CTTTCAAG 874  
DB 661 CTTTCAAG 692

RESULT 6  
LOCUS B0017479/c 629 bp mRNA linear EST 17-JUN-2002  
DEFINITION UI-H-DTI-awe-j-21-0-UI-s1 NCI CGAP DT1 Homo sapiens cDNA clone  
IMAGE:5888372.3', mRNA sequence.  
ACCESSION B0017479  
VERSION B0017479.1 GI:19752756  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILMIL at: <http://image.llnl.gov>  
The following repetitive elements were found in this cDNA  
Sequence: 458-519, >MIR#SINE/MIR  
Seq primer: M13 FORWARD  
POLYA=yes.

FEATURES  
source 1..629 Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5888372"  
/tissue\_type="Metastatic Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP DT1"  
/note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
NCI CGAP DT1 is a normalized cDNA library containing the  
following tissue(s): Metastatic Chondrosarcoma in lung.  
The library was constructed according to Bonaldo, Lemmon  
and Soares, Genome Research, 6:791-806, 1996. First strand  
cDNA synthesis was primed with an oligo-dT primer  
containing a Not I site. Double stranded cDNA was ligated  
to an EcoR I adaptor, digested with Not I, and cloned  
directionally into p773-Pac vector. The oligonucleotide  
used to prime the synthesis of first-strand cDNA contains  
a library tag sequence that is located between the Not I  
site and the (dT)18 tail. The sequence tag for this  
library is AACTGTTCCG.  
TAG\_TISSUE=lung metastatic chondrosarcoma  
TAG\_LIB=UI-H-DTI  
TAG\_SEQ=AACTGTTCCG"

ORIGIN  
Query Match 19.8%; Score 609; DB 12; Length 629;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2370 AAAGGCCCTGTATCAGCTCTTACAGATATTTGATTTTAACTCTTAACTA 2429  
DB 629 AAAGGCCCTGTATCAGCTCTTACAGATATTTGATTTTAACTCTTAACTA 570  
QY 2430 TTGAAGTTTAAAGGCGCTTAAAGAACTTAAGATCTTAAAGGCTTTCTGAGTTCA 2489  
DB 569 TTGAAGTTTAAAGGCGCTTAAAGAACTTAAGATCTTAAAGGCTTTCTGAGTTCA 510  
QY 2490 GAGAGGTTAAGTAACTTCTTCAAGGTACACAGCAAGTCTGTGGGTGACAGCAAGCT 2549  
DB 509 GAGAGGTTAAGTAACTTCTTCAAGGTACACAGCAAGTCTGTGGGTGACAGCAAGCT 450  
QY 2550 AGGCGTGGCATTTAGATACATACAGATGCTCCCTCTTATGTTGGCCCTGGG 2609  
DB 449 AGGCGTGGCATTTAGATACATACAGATGCTCCCTCTTATGTTGGCCCTGGG 390  
QY 2610 GCTTCAAGGCGCTTGGAGACATCTTGTCTCAACCTCTCCCTAGATCACTGTGAGGGT 2669  
DB 389 GCTTCAAGGCGCTTGGAGACATCTTGTCTCAACCTCTCCCTAGATCACTGTGAGGGT 330  
QY 2670 CCTGTAGATTTGTGTACACCATGCTCCATGTATATACAGTACACAGATGTACAC 2729  
DB 329 CCTGTAGATTTGTGTACACCATGCTCCATGTATATACAGTACACAGATGTACAC 270  
QY 2730 AGATGTACATGTTCCAGGCCCGCTCTGATACCTGACACCTGACCCGAGCTTGGCC 2789  
DB 269 AGATGTACATGTTCCAGGCCCGCTCTGATACCTGACACCTGACCCGAGCTTGGCC 210  
QY 2790 CCTGCTGCGCTGTGTCTCAAGGAGAGCTTCAACCTGCTGTCTCCCTTCCCAAC 2849  
DB 209 CCTGCTGCGCTGTGTCTCAAGGAGAGCTTCAACCTGCTGTCTCCCTTCCCAAC 150  
QY 2850 CACTGCTGAGCTTCTGAGAGACCAAGTACTTGTGCTGACCGGTGTGCGCCGCTC 2909  
DB 149 CACTGCTGAGCTTCTGAGAGACCAAGTACTTGTGCTGACCGGTGTGCGCCGCTC 90  
QY 2910 TCACCCAGGACACAGCCCGGACACATGATCTCCGTATACATTAATAAGTGGGT 2969  
DB 89 TCACCCAGGACACAGCCCGGACACATGATCTCCGTATACATTAATAAGTGGGT 30  
QY 2970 TGTACAAA 2978  
|||||||



Query Match  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

18.5%; Score 571; DB 13; Length 918;

ORIGIN  
 ino Ecori/XhoI sites using the following 5' adaptor:  
 GGACGACG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

QY	1017	GGATGCTGGGGCGGGGCTTTGGCCAGTGCCTTTACAGGAGGTGGCCCCAGATGCCACTGTG	1076
Db	1	GGATGCTGGGGCGGGGCTTTGGCCAGTGCCTTTACAGGAGGTGGCCCCAGATGCCACTGTG	60
QY	1077	CGCATCTCCCCACCGAAGGCCCGGCGACAGCTTTGCACAGACCTTAAGGACACAGACTG	1136
Db	61	CGCATCTCCCCACCGAAGGCCCGGCGACAGCTTTGTTCACAGACCTTAAGGACACAGACTG	120
QY	1137	GAGGCCCCCGGGCGCTGGCCCTCGCACATTTGTCGTGCTTCACAGCTTTCTGAAGTC	1196
Db	121	GAGGCCCCCGGGCGCTGGCCCTCGCACATTTGTCGTGCTTCACAGCTTTCTGAAGTC	180
QY	1197	GCTTGTCACAGCTCTTTGGTGGTTTATCATCTCTCTGTGGGAGGACATCTTGACGCC	1256
Db	181	GCTTGTCACAGCTCTTTGGTGGTTTATCTCTCTGTGGGAGGACATCTTGACGCC	240
QY	1257	TCAAGAGTTAGGACAGACTCAAGTTACACCTTCTCTCTGGGGTTGGAAGAATGTTG	1316
Db	241	TCAAGAGTTAGGACAGACTCAAGTTACACCTTCTCTCTGGGGTTGGAAGAATGTTG	300
QY	1317	ATGCGAAGGGGTGAGATTGCTGCGTCATATGGAGCTCTCTGGGACAAAGCTCAGAGTG	1376

QY	1377	AAAAGGACACAGAGGCGCATGTGAAAGGCTCTCTCTCTGGCATTAACCCAGCTT	1436
Db	361	AAAAGACACAGAAAGGCCAGATGAAAGGTCTCTCTCTCTGGCATTAACCCAGCTT	420
QY	1437	GGTTTGGGTGGCAGCTGGAGAAACTTCTCTCCAGCCCTGCACACTTAAAGCTCTGTTTC	1496
Db	421	GGTTTGGGTGGCAGCTGGAGAAACTTCTCTCTCCAGCCCTGCACACTTAAAGCTCTGTTTC	480
QY	1497	AGCTGCTCTTGCACATCCCTCTCCACATCCCATGACACACAAAGTTGGCCCCAGCTGGCC	1556
Db	481	AGCTGCTCTTGCACATCCCTCTCCACATCCCATGACACACAAAGTTGGCCCCAGCTGGCC	540
QY	1557	TGACATTGAGCCAGTGGACTGTGTCGTGA	1587
Db	541	TGACATTGAGCCAGTGGACTGTGTCGTGA	571
RESULT 10			
LOCUS	BF970943		
DEFINITION	BF970943	730 bp	mRNA
ACCESSION	602270226f1 NIH_MGC_84	Homo sapiens	linear
VERSION	BF970943		EST 22-JAN-2001
KEYWORDS	BF970943.1	GI:1238158	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE 1 (bases 1 to 730)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.



CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM9996 row: c column: 08  
High quality sequence stop: 688.

## FEATURES

source

1. 730

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:435835"

/tissue\_type="adrenal cortex carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_1ib="NIH\_MGC\_84"

/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site\_1:  
NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT  
primed. Average insert size 1.229 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 17.6%; Score 542; DB 10; Length 730;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 642; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

## FEATURES

source

1. 523

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2988716"

/tissue\_type="melanotic melanoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_1ib="NIH\_MGC\_20"

/note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Query Match 16.8%; Score 519; DB 10; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1952 CTTTCTTTTATGCTCCCATTTACAGAGATACCAAGAGCTGGAGAGGAGACGATG 2011  
241 CTTTCTTTTATGCTCCCATTTTACAGAGATACCAAGAGCTGGAGAGGAGACGATG 300  
2012 AGCACCAGGCCCCCAGCTTTGTCCTTACGAAATTCAGGGTACAGCTCCACTAGAACGAG 2071  
301 AGCACCAGGCCCCCAGCTTTGTCCTTACGAAATTCAGGGTACAGCTCCACTAGAACGAG 360  
2072 GGTGCGCTTACTAGTGTCTGCTGCTTCCCAACATTTATTAAGACCTTAAGGCTGGGTT 2131  
361 GGTGCGCTTACTAGTGTCTGCTGCTTCCCAACATTTATTAAGACCTTAAGGCTGGGTT 420  
2132 CACTGTGTCTAGAGAAACCAAGAGGCTCCCAAGTCTGTGCTCTGTGCGCGCTTGTGCTG 2191  
421 CACTGTGTCTAGAGAAACCAAGAGGCTCCCAAGTCTGTGCTCTGTGCGCGCTTGTGCTG 480  
2192 CCACCACTTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGAGCTGGAGACCAAGAAC 2251  
481 CCACCACTTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGAGCTGGAGACCAAGAAC 540  
2252 TGAGCTGGGAGAGGATCCGAGAGAAAGCTCAGGGGCGGGTCTTCTCCTTGTGCGCGGAT 2311  
541 TGAGCTGGGAGAGGATCCGAGAGAAAGCTCAGGGGCGGGTCTTCTCCTTGTGCGCGGAT 600  
2312 TGGGCTATGCTGGGTACACCATGTACTCAGGCATGTGTGGGTTT 2355  
601 TGGGCTATGCTGGGTACACCATGTACTCAGGCATGTGTGGGTTT 644

RESULT 11

BE275652 523 bp mRNA linear EST 13-JUL-2000  
LOCUS 601121184P1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:2988716 5',  
DEFINITION mRNA sequence.  
ACCESSION BE275652  
VERSION BE275652.1 GI:9150610  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC/DCTD/DRP  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLM9996 row: c column: 21  
High quality sequence stop: 521.

Location/Qualifiers  
1. 523  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2988716"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NIH\_MGC\_20"  
/note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

26 GCCGCGCCAGCCCGGCTCCGAGACGCGAGGAGAGATCCCGCGCACTGACCCGGG 85  
4 GCCGCGCCAGCCCGGCTCCGAGACGCGAGGAGAGATCCCGCGCACTGACCCGGG 63  
86 AGCCACCAAGACTCTGGAGAGCTCGGCGGCTGGAGACAGCAGGAGCTCCCGGAGCTCC 145  
64 AGCCACCAAGACTCTGGAGAGCTCGGCGGCTGGAGACAGCAGGAGCTCCCGGAGCTCC 123  
146 CGGCGCTTCCAGGAGCTCTTGAAGCGGTGCAAGAGCCCGGCCCATTTCCAGGCCCC 205  
124 CGGCGCTTCCAGGAGCTCTTGAAGCGGTGCAAGAGCCCGGCCCATTTCCAGGCCCC 183  
206 GAGCCATGATGAAGACTTTTGCAGCGGGAATCTCAGCTCAGGTGCGCGCAAAAACT 265  
184 GAGCCATGATGAAGACTTTTGCAGCGGGAATCTCAGCTCAGGTGCGCGCAAAAACT 243  
266 CATACCGCATGTGTGTGCTGCTCTCTGCGGTGGGCAAGAGCTCATGTGTCTGCT 325  
244 CATACCGCATGTGTGTGCTGCTCTCTGCGGTGGGCAAGAGCTCATGTGTCTGCT 303  
326 TCCCTCAATGCGCGCTTTGAGGAGCAGTACACACCAACATCAGAGATTCCACCGTAGG 385  
304 TCCCTCAATGCGCGCTTTGAGGAGCAGTACACACCAACATCAGAGATTCCACCGTAGG 363  
386 TATACCAATCCGCGGAGACATGTACAGCTCGACATCTGTGATACCTGTGGCAACCAACC 445



Db 364 TATACAAATCCGCGGCAGACATGATACAGCTCGATCCTGATACCTTCGACACACC 423  
Oy 446 CTTTCCCGCCATGCGAGGCTGTCCATCCTTCACAGGGGATGTCTTCAATCCGGTGTCA 505  
Db 424 CTTTCCCGCCATGCGAGGCTGTCCATCCTTCACAGGGGATGTCTTCAATCCGGTGTCA 483  
Oy 506 GCTTGATACCGGAGTCTTCGATGAGTCAAGCGCC 544  
Db 484 GCTTGATACCGGAGTCTTCGATGAGTCAAGCGCC 522

RESULT 12  
BE389944 623 bp mRNA linear EST 21-JUL-2000  
LOCUS 601282582P1 NIH\_MGC\_44 Homo sapiens CDNA clone IMAGE:3604196 5',  
DEFINITION mRNA sequence.  
ACCESSION BE389944  
VERSION BE389944.1 GI:9335309  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 623)  
NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L1CM250 row: d column: 21  
High quality sequence stop: 614.  
Location/Qualifiers  
1. .623  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3604196"  
/tissue\_type="endometrium, adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_44"  
/note="Organ: uterus; Vector: POTB7; Site 1: XhoI; Site 2:  
EcoRI; CDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN  
Query Match 16.7%; Score 516; DB 10; Length 623;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 96 GACTCTGGGAGGCTCGGCGGTGAGAGCAGCAGGAGCTCCCGAGCTCCCGGGCTTCC 155  
Oy 74 GACTCTGGGAGGCTCGGCGGTGAGAGCAGCAGGAGCTCCCGAGCTCCCGGGCTTCC 133  
Db 156 AGGAGAGCTCTGAGCCGTGCGCAGAGGCCCGCCGCGCATTCAGCCCGAGCGCATAT 215  
Oy 134 AGGAGAGCTCTGAGCCGTGCGCAGAGGCCCGCCGCGCATTCAGCCCGAGCGCATAT 193  
Db 216 GAAAGCTTTGTCCAGCGGGAAGTGCAGCGCTCAGTGTGCCCGCAAAAATCATACCGCAT 275  
Oy 194 GAAAGCTTTGTCCAGCGGGAAGTGCAGCGCTCAGTGTGCCCGCAAAAATCATACCGCAT 253

Oy 276 GGTGTGCTGGGAGCTCTTCGCGGTGGGCAAGAGCTCCATCGTCTCGCTTCAATGG 335  
Db 254 GGTGTGCTGGGAGCTCTTCGCGGTGGGCAAGAGCTCCATCGTCTCGCTTCAATGG 313  
Oy 336 CCGCTTTGAGGACCAATACACCCAGCATGAGAGCTTCCACCGTAAGTATACAAAT 395  
Db 314 CCGCTTTGAGGACCAATACACCCAGCATGAGAGCTTCCACCGTAAGTATACAAAT 373  
Oy 396 CCGGCGGACATGTACCAAGCTCGACATTCCTGATACCTCTGGCAACACCCCTCCCGC 455  
Db 374 CCGGCGGACATGTACCAAGCTCGACATTCCTGATACCTCTGGCAACACCCCTCCCGC 433  
Oy 456 CATGCGAGGCTGTCCATCTTCACAGGGGATGTCTTCATCTCGGTTCACCTGATTA 515  
Db 434 CATGCGAGGCTGTCCATCTTCACAGGGGATGTCTTCATCTCGGTTCACCTGATTA 493  
Oy 516 CCGGAGTCTTTCATGATGAGTCAAGCGCTTCAGAGCAGATCTTGAGGTCAAGTCTG 575  
Db 494 CCGGAGTCTTTCATGATGAGTCAAGCGCTTCAGAGCAGATCTTGAGGTCAAGTCTG 553  
Oy 576 CCGGAGTCTTTCATGATGAGTCAAGCGCTTCAGAGCAGATCTTGAGGTCAAGTCTG 611  
Db 554 CCGGAGTCTTTCATGATGAGTCAAGCGCTTCAGAGCAGATCTTGAGGTCAAGTCTG 589

RESULT 13  
BG119387 1185 bp mRNA linear EST 30-JAN-2001  
LOCUS BG119387  
DEFINITION 602349070P1 NIH\_MGC\_90 Homo sapiens CDNA clone IMAGE:4444168 5',  
LOCUS BG119387  
DEFINITION mRNA sequence.  
ACCESSION BG119387  
VERSION BG119387.1 GI:12612893  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1185)  
NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L1AM10219 row: k column: 17  
High quality sequence stop: 679.  
Location/Qualifiers  
1. 1185  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4444168"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_90"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC library."

ORIGIN  
Query Match 16.6%; Score 511; DB 12; Length 1185;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1930 TGGGGTGGGAGGCGCAGAGCTGCTTTTAAATGCTCCATTTACAGAGGATACCAACC 1989

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_id="70480820"
/clone_1db="SG0000215"
/note="Shotgun subclone library"

```

```
FEATURES
  source
    1..574
      /organism="Homo sapiens"
      /mol_type="mRNA"
```

```
/db_xref="taxon:9606"  
/clone="S22SNUI6n1-53-A09"  
/sex="F"  
/tissue_type="Ascites"  
/cell_type="Lymphoblast-like"  
/cell_line="SNU-16"  
/lab_host="DH10B"  
/clone_id="S22SNUI6n1"  
/note="Organ: Stomach; Vector: pT7T3-Pac; Site_1: EcoRI;  
Site_2: NotI; The S22SNUI6 library was contributed by the  
Soares laboratory and it was constructed as described by  
Bonaldi, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
Research 6(9): 791-806. RNA was prepared from harvested  
cells of SNU-16 culture. SNU-16 cell was obtained from  
Korean Cell Line Bank (KCLB). SNU-16 was established from  
ascitic fluids of Korean patients by Park U.G. et al.  
(1990), Cancer Res 50: 2773-2780."
```

## ORIGIN

```
Query Match      15.7%: Score 483; DB 12; Length 574;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1743 GGAATCATTTGTAAGAAAGCTAGTACTCCCTCTG3CCAGGCTCACCAGTT 1802  
D1 1 GGAATCATTTGTAAGAAAGCTAGTACTCCCTCTG3CCAGGCTCACCAGTT 60  
QY 1803 CAGATCCAGGCTCCACCCG3GAGCCTTCTCTGCTGCCAAACAGGATTCGCTG 1862  
D1 61 CAGATCCAGGCTCCACCCG3GAGCCTTCTCTGCTGCCAAACAGGATTCGCTG 120  
QY 1863 GCCTGTTGACATGAGATGACCTCGGCATGAGCTCCAGGTTTACAGACAATTGC 1922  
D1 121 GCCTGTTGACATGAGATGACCTCGGCATGAGCTCCAGGTTTACAGACAATTGC 180  
QY 1923 ACAAGCGTGGG3G3GAGG3CAGGAGCTGTTTTTTTAAATGCTCCATTTCACAGAGA 1982  
D1 181 ACAAGCGTGGG3G3GAGG3CAGGAGCTGTTTTTTTAAATGCTCCATTTCACAGAGA 240  
QY 1983 TACCAACGAGACTCGAGAG3G3ACAGATGAGACACAG3CCCACTTTGTCCTAGCAA 2042  
D1 241 TACCAACGAGACTCGAGAG3G3ACAGATGAGACACAG3CCCACTTTGTCCTAGCAA 300  
QY 2043 ATTGAGG3TACAGCTCCACTAGAACAGG3G3CCTTACTGCTGCTGCTTCTCAAGCA 2102  
D1 301 ATTGAGG3TACAGCTCCACTAGAACAGG3G3CCTTACTGCTGCTGCTTCTCAAGCA 360  
QY 2103 TTTATTAAGACCTTACTGGTGTGCTGCTTCACTGTCTTAGAAACCAAGAGG3TCCC 2162  
D1 361 TTTATTAAGACCTTACTGGTGTGCTGCTTCACTGTCTTAGAAACCAAGAGG3TCCC 420  
QY 2163 AGTCCTG3GCTCTG3CCG3CCTGCTG3CCCAACACTTCTGACACACAGG3G3G3GA 2222  
D1 421 AGTCCTG3GCTCTG3CCG3CCTGCTG3CCCAACACTTCTGACACACAGG3G3G3GA 480  
QY 2223 GGC 2225  
D1 481 GGC 483
```

Search completed: June 20, 2004, 17:38:12  
Job time : 7588 secs

THIS PAGE BLANK (USPTO)



PS Claim 4; Fig 1A; 73bp; English.

CC The present invention relates to a new polypeptide comprising a fully  
CC defined 266-amino acid sequence or its fragment, orthologue or allelic  
CC variant. The allelic variant or orthologue is encoded by a nucleic acid  
CC molecule that hybridizes under stringent conditions to the opposite  
CC strand of a nucleic acid molecule with a fully defined sequence of 3082  
CC or 11201 bp. The pharmaceutical composition comprising the agent that  
CC binds to the polypeptide is useful for treating a disease or condition  
CC mediated by a human Ras-like protein e.g. cancer. Likewise, the  
CC polypeptide is useful for identifying modulators of Ras-like protein  
CC activity or expression by screening a compound for the ability to  
CC stimulate or inhibit interaction between the Ras-like protein and an  
CC interacting molecule or Ras-like activity. The polypeptide is useful for  
CC diagnosing a disease, or predisposition to a disease, or treating  
CC diseases associated with an increase in apoptosis or cell proliferation  
CC e.g. AIDS (acquired immunodeficiency syndrome), Alzheimer's disease,  
CC Parkinson's disease, aplastic anaemia, allergies, atherosclerosis,  
CC osteoporosis or bacterial, viral, fungal, parasitic, protozoal or  
CC helminthic infections, and inflammation. The present nucleic acid  
CC sequence encodes the human Ras-like protein of the invention. This gene  
CC is located on chromosome 22

XX Sequence 3082 BP; 708 A; 962 C; 793 G; 619 T; 0 U; 0 Other;

Query Match 100.0%; Score 3082; DB 6; Length 3082;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 3082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCGTGGCGCGGGGAGAAAGCCGCGCCGCGGCTCCGAGCAGCCGCAAGG 60  
DB 1 GGGCGTGGCGCGGGGAGAAAGCCGCGCCGCGGCTCCGAGCAGCCGCAAGG 60  
QY 61 GAGGATCCCGCGGAGTGAACCCGGAAGCCACACAGACTTGGAGAGCTGGCGCTGA 120  
DB 61 GAGGATCCCGCGGAGTGAACCCGGAAGCCACACAGACTTGGAGAGCTGGCGCTGA 120  
QY 121 GCAGCAGGAGAGCTCCCGGAGCTCCCGGAGCTTCCAGGAGCTCTGAGCGGTCAG 180  
DB 121 GCAGCAGGAGAGCTCCCGGAGCTCCCGGAGCTTCCAGGAGCTCTGAGCGGTCAG 180  
QY 181 GGCCTGGCGCGGCGGATTCGAGCCGCGGATGAGAGCTTTGTCAGCGGGAATGCG 240  
DB 181 GGCCTGGCGCGGCGGATTCGAGCCGCGGATGAGAGCTTTGTCAGCGGGAATGCG 240  
QY 241 AGGCTCAATGTGCGCGGCAAAATCTATACCGCATGTGTGCTGTGGGTG 300  
DB 241 AGGCTCAATGTGCGCGGCAAAATCTATACCGCATGTGTGCTGTGGGTG 300  
QY 301 GGCAGAGAGCTCATGTGTGCTGCTCAATGAGCGGCTTGAAGACAGTACACAGC 360  
DB 301 GGCAGAGAGCTCATGTGTGCTGCTCAATGAGCGGCTTGAAGACAGTACACAGC 360  
QY 361 ACCATGAGAGCTTCCACCGTAAAGGTATACCATCGCGGAGCATGTATACAGCTGAC 420  
DB 361 ACCATGAGAGCTTCCACCGTAAAGGTATACCATCGCGGAGCATGTATACAGCTGAC 420  
QY 421 ATCCGATACCTCTGGAACCAACCCCTTCCCGGCAATGCGAGGCTGTCAATCTTCA 480  
DB 421 ATCCGATACCTCTGGAACCAACCCCTTCCCGGCAATGCGAGGCTGTCAATCTTCA 480  
QY 481 GGGGATGTCTTCAATCTGTGTGTTGAGCTGTAAACCGGAGTCTTGAATGAGTCAAG 540  
DB 481 GGGGATGTCTTCAATCTGTGTGTTGAGCTGTAAACCGGAGTCTTGAATGAGTCAAG 540  
QY 541 CGCTTGAAGACAGATCTTGAAGGTCAAGTCTCTGAGAAACAAAGAGAGCG 600  
DB 541 CGCTTGAAGACAGATCTTGAAGGTCAAGTCTCTGAGAAACAAAGAGAGCG 600  
QY 601 GGGAGAGTGGCAGATCTGTGTGGAACAAAGACGACGAGGAGCTGTGCGGAG 660  
DB 601 GGGAGAGTGGCAGATCTGTGTGGAACAAAGACGACGAGGAGCTGTGCGGAG 660

QY 661 GTGCCACCAACGAGAGCGGAGCTGTGTGTGCGGACGAGAACTCCGCTACTTGCAG 720  
DB 661 GTGCCACCAACGAGAGCGGAGCTGTGTGTGCGGACGAGAACTCCGCTACTTGCAG 720  
QY 721 GTGTGCGGCAAGAAACACCAACGATGATGTTCTACGATCTTCAAGATGAGC 780  
DB 721 GTGTGCGGCAAGAAACACCAACGATGATGTTCTACGATCTTCAAGATGAGC 780  
QY 781 AAGCTCCACACGAGATGAGCCCGCTGATGCGAAGATCTCCGTGAGTACGATGAC 840  
DB 781 AAGCTCCACACGAGATGAGCCCGCTGATGCGAAGATCTCCGTGAGTACGATGAC 840  
QY 841 GCTTCCACCCAGGAGCTTGTGATGCGCGGCTGAGAGATGAGAGCTTATGAGCATG 900  
DB 841 GCTTCCACCCAGGAGCTTGTGATGCGCGGCTGAGAGATGAGAGCTTATGAGCATG 900  
QY 901 GTTCGAGCTTGGCGCGCGGCGGCGGAGCTCAACAGTACCTCAAGTATCAAGGCGCA 960  
DB 901 GTTCGAGCTTGGCGCGCGGCGGCGGAGCTCAACAGTACCTCAAGTATCAAGGCGCA 960  
QY 961 GTCTTGGGAAAGCGCAGGCGCGGTGAGAGGCAAGTGCATTCAGTACGAGGAT 1020  
DB 961 GTCTTGGGAAAGCGCAGGCGCGGTGAGAGGCAAGTGCATTCAGTACGAGGAT 1020  
QY 1021 GGTGGGCGGCGGCTTGGCAGTGCCTTCAAGGAGGTGCGGCGGAGATGCGGCA 1080  
DB 1021 GGTGGGCGGCGGCTTGGCAGTGCCTTCAAGGAGGTGCGGCGGAGATGCGGCA 1080  
QY 1081 TCTCCCAACGAGGCGCGGCGGAGAGCTTGTTCACAGACTTATGAGCAGTACGAG 1140  
DB 1081 TCTCCCAACGAGGCGCGGCGGAGAGCTTGTTCACAGACTTATGAGCAGTACGAG 1140  
QY 1141 CCCCCGCGGCTGTGCTCCGCAATGTGTGCTTCTCAAGCTTCTGAGTCCGCTT 1200  
DB 1141 CCCCCGCGGCTGTGCTCCGCAATGTGTGCTTCTCAAGCTTCTGAGTCCGCTT 1200  
QY 1201 GTCCAGAGCTCTTGTGTGATTCATCTCTGTGTGAGGAGACATCTGACAGCTCAA 1260  
DB 1201 GTCCAGAGCTCTTGTGTGATTCATCTCTGTGTGAGGAGACATCTGACAGCTCAA 1260  
QY 1261 GAGTTAGGAGAGACTCAAGTTACCTTCTCTGTGGGTTGAAAGAAATGTTATGC 1320  
DB 1261 GAGTTAGGAGAGACTCAAGTTACCTTCTCTGTGGGTTGAAAGAAATGTTATGC 1320  
QY 1321 CAGAGGGGTGAGATGTGTGCTCATATGAGAGCTTCTGAGGAGACAGCTGAGTGA 1380  
DB 1321 CAGAGGGGTGAGATGTGTGCTCATATGAGAGCTTCTGAGGAGACAGCTGAGTGA 1380  
QY 1381 GAGACAGAGAGGCGAGATGAGAAAGTCTCTCTCTCTGAGCATTAACAGCTGAGT 1440  
DB 1381 GAGACAGAGAGGCGAGATGAGAAAGTCTCTCTCTCTGAGCATTAACAGCTGAGT 1440  
QY 1441 TGGGTGGAGCTGGAGAACTTCTCTCCAGCCCTGCAACTTTAGGCTGTGATGACT 1500  
DB 1441 TGGGTGGAGCTGGAGAACTTCTCTCCAGCCCTGCAACTTTAGGCTGTGATGACT 1500  
QY 1501 GCTCTGACACCCCTCCACCCCGAGACACACACAAAGTGGCCCCAGCTGGCGCTGAC 1560  
DB 1501 GCTCTGACACCCCTCCACCCCGAGACACACACAAAGTGGCCCCAGCTGGCGCTGAC 1560  
QY 1561 ATTGAGCAGTGAATCTGTGTGTAAGAGGAGGCTGAGCAACCTCTAGACAGGCCA 1620  
DB 1561 ATTGAGCAGTGAATCTGTGTGTAAGAGGAGGCTGAGCAACCTCTAGACAGGCCA 1620  
QY 1621 CCACTTGAACCAAGCCCACTCTGACCGGTTCTCAAGCTCTCTTGAAGTCTCTCC 1680  
DB 1621 CCACTTGAACCAAGCCCACTCTGACCGGTTCTCAAGCTCTCTTGAAGTCTCTCC 1680  
QY 1681 GCGGAGAGTGTGCTTGTGTGTGAGAGCTTGTGTGATGATGATGATGATGATGAT 1740  
DB 1681 GCGGAGAGTGTGCTTGTGTGTGAGAGCTTGTGTGATGATGATGATGATGATGAT 1740  
QY 1741 ATGGAATCAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800







QY 2126 TGGGTTCACTGTGTCTTAGAAGAACAGAGGGTCCCGAGTCTTGACCTTCCGCGCCCT 2185  
 DB 2123 TGGGTTCACTGTGTCTTAGAAGAACAGAGGGTCCCGAGTCTTGACCTTCCGCGCCCT 2182  
 QY 2186 GCTGCCCCCACCACCTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGCAGCTGGAGCC 2245  
 DB 2183 GCTGCCCCCACCACCTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGCAGCTGGAGCC 2242  
 QY 2246 CAGAACTGAGCTTGGAGAGGATCCGACAGAAAAGCTCAGGGCGGGTCTTCTCTTGGGCC 2305  
 DB 2243 CAGAACTGAGCTTGGAGAGGATCCGACAGAAAAGCTCAGGGCGGGTCTTCTCTTGGGCC 2302  
 QY 2306 CGGGATTGGGCTATGCTGGGATACACCATGTACTCAGGCGATGGGTTTGAACCCATA 2365  
 DB 2303 CGGGATTGGGCTATGCTGGGATACACCATGTACTCAGGCGATGGGTTTGAACCCATA 2362  
 QY 2366 AACCAAGGCGCTTGTGATCAGCTCTTAAACAATATTTTATTTTAACTCTCTTAA 2425  
 DB 2363 AACCAAGGCGCTTGTGATCAGCTCTTAAACAATATTTTATTTTAACTCTCTTAA 2422  
 QY 2426 CATATTGAAGTTTAAAGGCGCTTAAAGAACTTATGATCTTCTATTTGGGCTTTTCTGAGG 2485  
 DB 2423 CATATTGAAGTTTAAAGGCGCTTAAAGAACTTATGATCTTCTATTTGGGCTTTTCTGAGG 2482  
 QY 2486 TTTCAGAGAGGTAGTAACTTCCCTCCAGGTACACAGAAATCTGTGGTGGAGAAAGCA 2545  
 DB 2483 TTTCAGAGAGGTAGTAACTTCCCTCCAGGTACACAGAAATCTGTGGTGGAGAAAGCA 2542  
 QY 2546 AGCTAGGCGCTGGGATTCAGTACATACACAGATGTCCTCTCTTGTAGTCTTGCGCC 2605  
 DB 2543 AGCTAGGCGCTGGGATTCAGTACATACACAGATGTCCTCTCTTGTAGTCTTGCGCC 2602  
 QY 2606 TGGGCGCTTCAAGGCGCTTGGGACATCTTGTCTTCAACCTTCTCCCTAGATCAGTCTGTA 2665  
 DB 2603 TGGGCGCTTCAAGGCGCTTGGGACATCTTGTCTTCAACCTTCTCCCTAGATCAGTCTGTA 2662  
 QY 2666 GGGTCCCTGATGATTTGTGTACACCAAGTGTATATTAACAAGTACACACAGATGTA 2725  
 DB 2663 GGGTCCCTGATGATTTGTGTACACCAAGTGTATATTAACAAGTGTATATTAACAAGTGTAC 2722  
 QY 2726 ACACAGATGTACATGATCTCCAGGCGCTCTGATACCTGACCTGACCCGACCTT 2785  
 DB 2723 ACACAGATGTACATGATCTCCAGGCGCTCTGATACCTGATACCTGACCTGACCCGACCTT 2782  
 QY 2786 GGGCGCTGCTGCTGTGTGTCTCAAGAGCAGCTCCAACTGCTGTCTCTCTCCCTTCC 2845  
 DB 2783 GGGCGCTGCTGCTGTGTGTCTCAAGAGCAGCTCCAACTGCTGTCTCTCTCTCCCTTCC 2842  
 QY 2846 CACCACTGCTGAGCGCTTCTGAGAGACCAAGTACTTGGCTGACACCGGTGTGGGCC 2905  
 DB 2843 CACCACTGCTGAGCGCTTCTGAGAGACCAAGTACTTGGCTGACACCGGTGTGGGCC 2902  
 QY 2906 GCTCTACCCAGAGCAGGCGCGCCACCATGATCTCGGTACACTATCAATAAAGTG 2965  
 DB 2903 GCTCTACCCAGAGCAGGCGCGCCACCATGATCTCGGTACACTATCAATAAAGTG 2962  
 QY 2966 GGTGTTGTTACA 2976  
 DB 2963 GGTGTTGTTACA 2973

RESULT 3  
 ABX72012  
 ID ABX72012 standard; DNA; 2973 BP.

XX ABX72012;  
 DT 12-MAR-2003 (first entry)  
 XX DNA encoding human tumour endothelial marker TEM 2.  
 XX Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;

KW Tumour endothelial marker; normal endothelial marker; PEM;  
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
 KW neovascularization; immune response; cytostatic; antidiabetic; gene;  
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic; ds.  
 OS Homo sapiens.  
 PN MO200283874-A2.  
 XX 24-OCT-2002.  
 PD 10-APR-2002; 2002MO-US008253.  
 PF 11-APR-2001; 2001US-0282850P.  
 PR 06-FEB-2002; 2002US-0354262P.  
 XX (UYUO ) UNIV JOHNS HOPKINS.  
 PA Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;  
 PI WPI; 2003-093016/08.  
 DR P-PSDB; AB054440.  
 XX New purified human transmembrane protein, designated as tumor endothelial  
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,  
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.  
 XX Disclousure; Page 154-155; 374pp; English.  
 PS The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumours as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neovascularization or tumour angiogenesis, for  
 CC inducing an immune response to tumour endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumours. The present  
 CC sequence represents a human TEM or NEM gene of the invention  
 XX  
 SO Sequence 2973 BP; 596 A; 967 C; 787 G; 623 T; 0 U; 0 Other;  
 Query Match 94.1%; Score 2900; DB 7; Length 2973;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 26 GCGCGCGCCAGCGCGGTCTCCGAGCAGCGGAGAGATCCCGCGCACTGACCCGGG 85  
 DB 23 GCGCGCGCCAGCGCGGTCTCCGAGCAGCGGAGAGATCCCGCGCACTGACCCGGG 82  
 QY 86 AGCCACCAAGACTCTGGAGAGGCTCGGCGGCTGGAGAGCAGCAGCTCCCGGAGCTCC 145  
 DB 83 AGCCACCAAGACTCTGGAGAGGCTCGGCGGCTGGAGAGCAGCAGCTCCCGGAGCTCC 142  
 QY 146 CGGCGCTTCCAGGAGCGCTCTGAGCGCGTGCAGAGGCGCGCGCCATTTCCAGGCC 205  
 DB 143 CGGCGCTTCCAGGAGCGCTCTGAGCGCGTGCAGAGGCGCGCGCCATTTCCAGGCC 202  
 QY 206 GAGCCATGATGAAACATTTTGTCCAGCGGAACTGCAAGTGTGCGCGCCCAAAACT 265  
 DB 203 GAGCCATGATGAAACATTTTGTCCAGCGGAACTGCAAGTGTGCGCGCCCAAAACT 262  
 QY 266 CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325  
 DB 263 CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 322  
 QY 326 TCTTCAATGGCGGCTTTGAGAGCAGTACACACCACCATGAGAGACTTCCACGGTAAG 385  
 DB 323 TCTTCAATGGCGGCTTTGAGAGCAGTACACACCACCATGAGAGACTTCCACGGTAAG 382

QY 386 TATACAACTCCGGGGGCAATGATACCACTGCACTCTGGAATACCTTGGCAACACC 445  
Db 383 TATACAACTCCGGGGGCAATGATACCACTGCACTCTGGAATACCTTGGCAACACC 442  
QY 446 CCTTCCCGCCAGTGGCGCAGCTGTCCATCTCAAGGGATGTCTTCATCTGTGTTC 505  
Db 443 CCTTCCCGCCAGTGGCGCAGCTGTCCATCTCAAGGGATGTCTTCATCTGTGTTC 502  
QY 506 GCCTGGATTAACCGGAGTCTTGTGATGAGTCAAGCGCTTCAAGAGCATCTGGAG 565  
Db 503 GCCTGGATTAACCGGAGTCTTGTGATGAGTCAAGCGCTTCAAGAGCATCTGGAG 562  
QY 566 TCAAGCTCCGCTGAAGAACCAAGAACGCGCGGAGCTGCCCAAGTCACTGTG 625  
Db 563 TCAAGCTCCGCTGAAGAACCAAGAACGCGCGGAGCTGCCCAAGTCACTGTG 622  
QY 626 GCAACAGAACCAACGCGAGCTGTGCGCCAGTGGCCCAACCGAGGCCAGCTGC 685  
Db 623 GCAACAGAACCAACGCGAGCTGTGCGCCAGTGGCCCAACCGAGGCCAGCTGC 682  
QY 686 TGGTGTCCGGGCGACGAGAACTCCGCTTACTTGAAGTGTGCGCCAGAAAGAACCAACG 745  
Db 683 TGGTGTCCGGGCGACGAGAACTCCGCTTACTTGAAGTGTGCGCCAGAAAGAACCAACG 742  
QY 746 TGGAGAGATGTTCTACGTCCTTACGATGAGGCAAGTGCACACAGATGAGGCCCC 805  
Db 743 TGGAGAGATGTTCTACGTCCTTACGATGAGGCAAGTGCACACAGATGAGGCCCC 802  
QY 806 CCTTCATTCGAAGATCTCCGTGAGTACGATGAGCGCTTCCACCCAGGCCCTTCTGA 865  
Db 803 CCTTCATTCGAAGATCTCCGTGAGTACGATGAGCGCTTCCACCCAGGCCCTTCTGA 862  
QY 866 TGGCGCGGTCAAGAGATGAGCGCTTATGAGTGTCTGCGCTTCCGCGCGCGCCCA 925  
Db 863 TGGCGCGGTCAAGAGATGAGCGCTTATGAGTGTCTGCGCTTCCGCGCGCGCCCA 922  
QY 926 GCGTCAACAGTACCTCAAGTACATCAAGGCCCAAGGTCTTCCGGGAGGCCAGGCCG 985  
Db 923 GCGTCAACAGTACCTCAAGTACATCAAGGCCCAAGGTCTTCCGGGAGGCCAGGCCG 982  
QY 986 AAGAGGCAAGTGCACATCCAGTACGAGCGAGGATGTGAGGCGGGGCTTGGCCAGTGC 1045  
Db 983 AAGAGGCAAGTGCACATCCAGTACGAGCGAGGATGTGAGGCGGGGCTTGGCCAGTGC 1042  
QY 1046 TTCAGGGAGTGGGCCCCAGATGCCCATGTGGGCAATCTCCCAACCGAGGCCCGGACAG 1105  
Db 1043 TTCAGGGAGTGGGCCCCAGATGCCCATGTGGGCAATCTCCCAACCGAGGCCCGGACAG 1102  
QY 1106 GTCTTGTTCACAGACTTAAAGGACCACTGAGAGGCCCGCGGCGCTGGCCCTCCGCAAT 1165  
Db 1103 GTCTTGTTCACAGACTTAAAGGACCACTGAGAGGCCCGCGGCGCTGGCCCTCCGCAAT 1162  
QY 1166 TCGTGTGCTTCTCAAGACTTTCGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCAT 1225  
Db 1163 TCGTGTGCTTCTCAAGACTTTCGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCAT 1222  
QY 1226 CTCCTCTGAGGAGGACATCTGACGCTCAAGATTTAGGAGGAGACTCAATTACA 1285  
Db 1223 CTCCTCTGAGGAGGACATCTGACGCTCAAGATTTAGGAGGAGACTCAATTACA 1282  
QY 1286 CCTTCTCTCTGAGGAGGAGAAATGTGATGACAGAGGGGTGAGATTGCTGCTCA 1345  
Db 1283 CCTTCTCTCTGAGGAGGAGAAATGTGATGACAGAGGGGTGAGATTGCTGCTCA 1342  
QY 1346 TATGAGGCTCTTGGGACCAAGCTTGAAGTGAAGAAAGAACAGAGGCCAGATGAAG 1405  
Db 1343 TATGAGGCTCTTGGGACCAAGCTTGAAGTGAAGAAAGAACAGAGGCCAGATGAAG 1402  
QY 1406 GTCTCTCTCTCTGAGGAGTAAACCCAGCTTGTGTGTGGGAGCTGGGAAATTTCTC 1465  
Db 1403 GTCTCTCTCTCTGAGGAGTAAACCCAGCTTGTGTGTGGGAGCTGGGAAATTTCTC 1462

QY 1466 TCCAGCCCTGCAACTCTTAAGCTGTGAGTCAAGTGCCTTGCACCCCTCCACCCCA 1525  
Db 1463 TCCAGCCCTGCAACTCTTAAGCTGTGAGTCAAGTGCCTTGCACCCCTCCACCCCA 1522  
QY 1526 GCACACACAAAGTTGGGCCCCAGCTGCGCTGACATTTAGAGCAGATGACTGTGTCTG 1585  
Db 1523 GCACACACAAAGTTGGGCCCCAGCTGCGCTGACATTTAGAGCAGATGACTGTGTCTG 1582  
QY 1586 AAGGGGGGTGGCCACACTCTTAGACACGCGCCACCTTAGACCAACCCCACTGCTG 1645  
Db 1583 AAGGGGGGTGGCCACACTCTTAGACACGCGCCACCTTAGACCAACCCCACTGCTG 1642  
QY 1646 ACCGCTTCTAGCTCTCTCTAGGATCCCTCCGCGGACAGTGTGCTTGTGTG 1705  
Db 1643 ACCGCTTCTAGCTCTCTCTAGGATCCCTCCGCGGACAGTGTGCTTGTGTG 1702  
QY 1706 TTGCACTGTTTTGTCATGATAGTAAAGTAAATGAATCAATGATCTGTAAAGC 1765  
Db 1703 TTGCACTGTTTTGTCATGATAGTAAAGTAAATGAATCAATGATCTGTAAAGC 1762  
QY 1766 CTAGTACTCCCTCTTGGCCAGGCTTACCCAGTTCAGATCCAGGCTTCCACCGGG 1825  
Db 1763 CTAGTACTCCCTCTTGGCCAGGCTTACCCAGTTCAGATCCAGGCTTCCACCGGG 1822  
QY 1826 ACGCTTCTCTGCTGCTCCGAAACAGGGTTTCGGTGGCTGTTGCACTGATGA 1885  
Db 1823 ACGCTTCTCTGCTGCTCCGAAACAGGGTTTCGGTGGCTGTTGCACTGATGA 1882  
QY 1886 CCTCCGCAATGAGCTCCACGGTTTACAGAAATTTGCAAGCGTGGGTGGGAGGCA 1945  
Db 1883 CCTCCGCAATGAGCTCCACGGTTTACAGAAATTTGCAAGCGTGGGTGGGAGGCA 1942  
QY 1946 GACACTGTTTTTTTAAAGTCTCCATTTCAAGAGATACACCGAGCTGAGGGGAC 2005  
Db 1943 GACACTGTTTTTTTAAAGTCTCCATTTCAAGAGATACACCGAGCTGAGGGGAC 2002  
QY 2006 ACGATGAGCACAGGCCCCCACTTGTCCCTAGCAAAATTCAGGGTACAGCTCCACCTAG 2065  
Db 2003 ACGATGAGCACAGGCCCCCACTTGTGTCCCTAGCAAAATTCAGGGTACAGCTCCACCTAG 2062  
QY 2066 AACCAAGTGCCTTCTACTGTGCTGCTTCAAGCAATTTTAAAGCACTTACTGAGTGC 2125  
Db 2063 AACCAAGTGCCTTCTACTGTGCTGCTTCAAGCAATTTTAAAGCACTTACTGAGTGC 2122  
QY 2126 TGGGTTCACTGTGTCTTAAGAAACCAAGAGGTTCCCAAGTCTGAGCTTCCGCGCT 2185  
Db 2123 TGGGTTCACTGTGTCTTAAGAAACCAAGAGGTTCCCAAGTCTGAGCTTCCGCGCT 2182  
QY 2186 GCTGCCCAACCACTTCTGACACACAGCGGTGGGAGGCGGGAGGAGAGCTGGAGCC 2245  
Db 2183 GCTGCCCAACCACTTCTGACACACAGCGGTGGGAGGCGGGAGGAGAGCTGGAGCC 2242  
QY 2246 CAGAACTGAGCTTGGAGGAGTCCGACAGAAAGCTCAGGCGGAGTCTTCTGTGTG 2305  
Db 2243 CAGAACTGAGCTTGGAGGAGTCCGACAGAAAGCTCAGGCGGAGTCTTCTGTGTG 2302  
QY 2306 CGGATTTGGGCTTATGCTGGGATACCACTGATCTCAGGCAATGGGTTTGAACCAATA 2365  
Db 2303 CGGATTTGGGCTTATGCTGGGATACCACTGATCTCAGGCAATGGGTTTGAACCAATA 2362  
QY 2366 AACCAAGGCCCTTGTATCAGCTCTTAAAGATATTTTGTATTTTAAATCTCTTAA 2425  
Db 2363 AACCAAGGCCCTTGTATCAGCTCTTAAAGATATTTTGTATTTTAAATCTCTTAA 2422  
QY 2426 CATATTGAAGTTTAAAGGCCCTTAAGAACTTATGATCTTCTATTTGGGCTTTCTAG 2485  
Db 2423 CATATTGAAGTTTAAAGGCCCTTAAGAACTTATGATCTTCTATTTGGGCTTTCTAG 2482  
QY 2486 TTCAAGAGGGTATGATCTTCTCAGGTCACACACAGTCTGTGGGTGGGAGAGCA 2545  
Db 2483 TTCAAGAGGGTATGATCTTCTCAGGTCACACACAGTCTGTGGGTGGGAGAGCA 2542  
QY 2546 AGCTAGGCTGGGCAATTCAGTACATACAGATGTGCTCTCTCTTGTATGTGGCCCC 2605

Db	2543	AGCTAGCCCTGGGCACTTAGTACATACACAGATGTGCTCCTCTCTTGATGCTTGGCCCC	2602
Qy	2606	TGGGGCCCTTCAGGGCTTTGGAGCATCTTGTCTCAACCCCTCCCTAGTCAGTCTGGA	2665
Db	2603	TGGGGCCCTTCAGGGCTTTGGGACATCTGTCTCAACCTCTCCCTTAATCAGTCTGGA	2662
Qy	2666	GAGTCCCTGTAGATATTGTGTATACACCATGCCCATGTATATACAGTACACACAGATGTAC	2725
Db	2663	GGGTCCCTGTAGATATTGTGTATACACCATGCCCATGTATATACAGTACACACAGATGTAC	2722
Qy	2726	ACACAGATGTACACATGCTCCAGCCCCCAGCTCTGATACCTTGACCTTGACACCCAGCCTT	2785
Db	2723	ACACAGATGTACACATGCTCCAGCCCCCAGCTCTGATACCTTGACACCTTGACACCCAGCCTT	2782
Qy	2786	GGCCCTCTCCGCGCTGTGTGCTCAAAAGAGAGAGCTCAACCCCTGCTGTCCCTTCC	2845
Db	2783	GGCCCTCTCTGCGCTGTGTGCTCAAAAGAGAGAGCTCAACCCCTGCTGTGTCCCTTCC	2842
Qy	2846	CACCCACTGCTGAGCCTTCTTGAGCAGACCAAGTACCTTGGCTGCAACCGATGTGAGCC	2905
Db	2843	CACCCACTGCTGAGCCTTCTTGAGCAGACCAAGTACCTTGGCTGCAACCGATGTGTGAGCC	2902
Qy	2906	GCTCTCACCCAGGCAACAGCCCCCGCACCATGATCTCCGTGTACATATCAATAAAGTG	2965
Db	2903	GCTCTCACCCAGGCAACAGCCCCCGCACCATGATCTCCGTGTACATATCAATAAAGTG	2962
Qy	2966	GGTTTGTGTACA 2976	
Db	2963	GGTTTGTGTACA 2973	

RESULT 4	
ABL92076	
ID	ABL92076 standard; cDNA; 2832 BP.
XX	
AC	ABL92076;
XX	
DT	30-MAY-2002 (first entry)
XX	
DE	Human Tumour Endothelial Marker polynucleotide SEQ ID NO 174.
XX	
KW	Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
KW	normal endothelial marker; pan-endothelial marker; immunostimulant;
KW	antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
KW	polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW	psoriasis; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200210217-A2.
XX	
PD	07-FEB-2002.
XX	
PF	01-AUG-2001; 2001WO-US024031.
XX	
PR	02-AUG-2000; 2000US-0222599P.
PR	11-AUG-2000; 2000US-0224360P.
PR	11-APR-2001; 2001US-0282850P.
XX	
PA	(UYJO ) UNIV JOHNS HOPKINS.
XX	
P1	St Croix B, Kinzler KW, Vogelstein B;
XX	
DR	WPI; 2002-291856/33.
XX	
PT	An isolated molecule comprising an antibody variable region which
PT	specifically binds to an extracellular domain of a tumor endothelial
PT	marker (TEM) protein, useful for inhibiting tumor growth.
XX	
PS	Disclosure; Page 119-120; 33pp; English.
CC	The invention relates to an isolated molecule comprising an antibody

CC variable region which specifically binds to an extracellular domain of a  
CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,  
CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM  
CC proteins have cytostatic, immunostimulant and antiangiogenic activity.  
CC They are useful for inhibiting tumour growth, neoplasia and angiogenesis in subjects  
CC bearing a vascularised tumour, polycystic kidney disease, diabetic  
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM  
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)  
CC are disclosed, as are marker oligonucleotide sequences: tumour  
CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal  
CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers  
CC (PEM) ABL91903-ABL91995  
CC  
CC  
CC Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 U; 0 Other;

Query Match	Similarity	Score	DB	Length
89.5%;	100.0%;	2757;	DB 6;	2832;
Best Local Match	Conservative	Mismatches	Indels	Gaps
2807;	0;	1;	0;	0;
199	CAGCCCGAGCCATGATGAAAGACTTTTGTCCAGCGGGAACTTGACGCTCAGTGTGCCGCC	258		
25	CAGCCCGAGCCATGATGAAAGACTTTTGTCCAGCGGGAACTTGACGCTCAGTGTGCCGCC	84		
259	AAAAAATCTCATCCGCATGTGTGTCTGGGTGCTCTCTGGGTGGGCAAGAGTCCATGTG	318		
85	AAAAAATCTCATCCGCATGTGTGTCTGGGTGCTCTCTGGGTGGGCAAGAGTCCATGTG	144		
319	TTCTGCTTCCCATATGGCCGCTTTGAGGACCAAGTACACCCACATCGAGGACTTCCAC	378		
145	TTCTGCTTCCCATATGGCCGCTTTGAGGACCAAGTACACCCACATCGAGGACTTCCAC	204		
379	CGTAAGGTATACCAATCCGCGGCGAGATGTACCAGCTCGACATCTTGATCTTGAC	438		
205	CGTAAGGTATACCAATCCGCGGCGAGATGTACCAGCTCGACATCTTGATCTTGAC	264		
439	AACCAACCCCTTCCCGCCCATGCGAGGCTGTCCATCTCACAGGGGATGTCTTATCTGT	498		
265	AACCAACCCCTTCCCGCCCATGCGAGGCTGTCCATCTCACAGGGGATGTCTTATCTGT	324		
499	GTGTTCAGCCCTGATTAACCGGGAACTCCCTGATATAGGTCAAGCGCTTCAGAACAGATC	558		
325	GTGTTCAGCCCTGATTAACCGGGAACTCCCTGATATAGGTCAAGCGCTTCAGAACAGATC	384		
559	CTGAGAGGTCAAGTCTCTGCTGAAAGAACAGACCAAGGAGGCGCGAGCTGCCATGATC	618		
385	CTGAGAGGTCAAGTCTCTGCTGAAAGAACAGACCAAGGAGGCGCGAGCTGCCATGATC	444		
619	ATCTGTGGCAACAAAGAACACACGCGAGCTGTGCCGCAAGTGGCCACACCGAGGCC	678		
445	ATCTGTGGCAACAAAGAACACACGCGAGCTGTGCCGCAAGTGGCCACACCGAGGCC	504		
679	GAGCTGCTGCTGTGGGCGAGACGAAACTCCGCGCTACTTCAGAGTGTGGGCCAAGAAGAC	738		
505	GAGCTGCTGCTGTGGGCGAGACGAAACTCCGCGCTACTTCAGAGTGTGGGCCAAGAAGAC	564		
739	ACCAACGTGACGAGATGTTCTAAGTCTCTTTCAGATGGCCAAAGCTTGCCACAGATG	798		
565	ACCAACGTGACGAGATGTTCTAAGTCTCTTTCAGATGGCCAAAGCTTGCCACAGATG	624		
799	AGCCCGCGCTGCAATCGCAAGATTTCCGTGCAATGACGCTTGACCCCAAGGCC	858		
625	AGCCCGCGCTGCAATCGCAAGATTTCCGTGCAATGACGCTTGACCCCAAGGCC	684		
859	TTCTGCAATGGCGCGGTGCAAGGAAATGGAAGCCATGGACATGCTGGCCCTTGCGCCGC	918		
685	TTCTGCAATGGCGCGGTGCAAGGAAATGGAAGCCATGGACATGCTGGCCCTTGCGCCGC	744		
919	CGCCCAAGCTCAACAGTGAATCTTCAAGTATCATCAAGGCCAAGATCTTTCGGGAAGCGAG	978		
745	CGCCCAAGCTCAACAGTGAATCTTCAAGTATCATCAAGGCCAAGATCTTTCGGGAAGCGAG	804		
979	GCCCTGAGAGGACCAAGTGAATCTTCAAGTGAACCAAGGATGTTGGGCGCGGCTTGCC	1038		

Db 805 GCCCGTGAAGAGGACAGATGCAATCCAGTAGAGGAGATGCTGGGCGGGCTTGGC 864  
Qy 1039 CAGTCCCTTCAAGGAGTGGCCCCAGATGCCCACTGTGGCATCTCCCAACCGAGGCCCC 1098  
Db 865 CAGTCCCTTCAAGGAGTGGCCCCAGATGCCCACTGTGGCATCTCCCAACCGAGGCCCC 924  
Qy 1099 GGCACAGCTCTGTTCAACAGACCTTAAGGACAGACTGGAGGCCGCCCGGGCGCTGGCTC 1158  
Db 925 GGCACAGCTCTGTTCAACAGACCTTAAGGACAGACTGGAGGCCGCCCGGGCGCTGGCTC 984  
Qy 1159 GGCACATGCTGCTGCTCTCAACAGCTTCTGAGTCCGCTTGTCCACAGCTCTTGGTG 1218  
Db 985 GGCACATGCTGCTGCTCTCAACAGCTTCTGAGTCCGCTTGTCCACAGCTCTTGGTG 1044  
Qy 1219 GTTTCATCTCTCTGAGGAGGACATCTGCAAGCCCTCAGAGTTGAGGAGAGATCTCA 1278  
Db 1045 GTTTCATCTCTCTGAGGAGGACATCTGCAAGCCCTCAGAGTTGAGGAGAGATCTCA 1104  
Qy 1279 AGTTACACCTTCTCTCTGAGGAGGAGATGTTGATGTCAGAGGAGGAGATGTC 1338  
Db 1105 AGTTACACCTTCTCTCTGAGGAGGAGATGTTGATGTCAGAGGAGGAGATGTC 1164  
Qy 1339 TGCGTCATATGAGACCTCTCTGGGACAGCTCTGAGATGAAAGACACAGAGGCCAGAT 1398  
Db 1165 TGCGTCATATGAGACCTCTCTGGGACAGCTCTGAGATGAAAGACACAGAGGCCAGAT 1224  
Qy 1399 GAGAAAGGCTCTCTCTCTGAGGAGGAGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1458  
Db 1225 GAGAAAGGCTCTCTCTCTGAGGAGGAGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1284  
Qy 1459 ACTTCTCTCCAGCCCTGCAACTCTTAAGCTCTGAGTCCCTGCAACCCCTCC 1518  
Db 1285 ACTTCTCTCCAGCCCTGCAACTCTTAAGCTCTGAGTCCCTGCAACCCCTCC 1344  
Qy 1519 ACCCCACAGACACACAGATGGGCCCGGAGTGGCCCTGAGATGAGCGAGTCT 1578  
Db 1345 ACCCCACAGACACACAGATGGGCCCGGAGTGGCCCTGAGATGAGCGAGTCT 1404  
Qy 1579 GTGTCTGAAGGGGGGCTGAGCCCACTCTCTAGACAGAGCCCACTTAAGACAGGCCCA 1638  
Db 1405 GTGTCTGAAGGGGGGCTGAGCCCACTCTCTAGACAGAGCCCACTTAAGACAGGCCCA 1464  
Qy 1639 CCTCTGACCGGCTCTCAAGCTCTCTCTAGAGTCCCTCGGCCGACAGTGTGCTTT 1698  
Db 1465 CCTCTGACCGGCTCTCAAGCTCTCTCTAGAGTCCCTCGGCCGACAGTGTGCTTT 1524  
Qy 1699 GTTGTGTGTGACAGCTGTTTGTGTGATGATGATGATGATGATGATGATGATGATGATG 1758  
Db 1525 GTTGTGTGTGACAGCTGTTTGTGTGATGATGATGATGATGATGATGATGATGATGATG 1584  
Qy 1759 TAAAGCCTAGTGAATCCT 1818  
Db 1585 TAAAGCCTAGTGAATCCT 1644  
Qy 1819 ACCCGGAGCGCT 1878  
Db 1645 ACCCGGAGCGCT 1704  
Qy 1879 ACATTGACCTCGCATGAGCTCCAGGTTTACAGAAATTGACACAGCGCTGGGTGG 1938  
Db 1705 ACATTGACCTCGCATGAGCTCCAGGTTTACAGAAATTGACACAGCGCTGGGTGG 1764  
Qy 1939 CAGGCGCAGACTGCTTTTAAATGCTCCATTTTCAAGAGATACCAAGGAGCTGG 1998  
Db 1765 CAGGCGCAGACTGCTTTTAAATGCTCCATTTTCAAGAGATACCAAGGAGCTGG 1824  
Qy 1999 AGGGGACAGATGAGACACAGGCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2058  
Db 1825 AGGGGACAGATGAGACACAGGCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1884  
Qy 2059 CACCTAGAACAAGGCTGCTCTACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2118  
Db 1885 CACCTAGAACAAGGCTGCTCTACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1944

Qy 2119 TGGGCTGGGTTTCACTGTCTTATGAAACCAAGGGTCCCAAGTCTGGCTCTGCTC 2178  
Db 1945 TGGGCTGGGTTTCACTGTGTCTTAAAGAACCAAGGGTCCCAAGTCTGGCTCTGCTC 2004  
Qy 2179 GCGCCCTGCTGCCCCACACCTTCTGACACACAGCGGTGGGGAGGCGGGAGGAGCAGC 2238  
Db 2005 GCGCCCTGCTGCCCCACACCTTCTGACACACAGCGGTGGGGAGGCGGGAGGAGCAGC 2064  
Qy 2239 TGGGACCCAGAACTGAGCTTGGAGGAGTCCGACAGAAAGCTCAGGGCGGCTCTGCTC 2298  
Db 2065 TGGGACCCAGAACTGAGCTTGGAGGAGTCCGACAGAAAGCTCAGGGCGGCTCTGCTC 2124  
Qy 2239 TTTGGCCCGGAGATTTGGGTTATGCTGGGAGTCCAGCATGATCTCAGGATGAGGTTTGA 2358  
Db 2125 TTTGGCCCGGAGATTTGGGTTATGCTGGGAGTCCAGCATGATCTCAGGATGAGGTTTGA 2184  
Qy 2359 ACCCATAAACCAAGGCGCTTGTCACTGACTCTTAAACAGATATTTTGTATTTTAACT 2418  
Db 2185 ACCCATAAACCAAGGCGCTTGTCACTGACTCTTAAACAGATATTTTGTATTTTAACT 2244  
Qy 2419 CTCTAAACATATTTGAATTTAGGGCCCTTAAGAACTTATGATCTTATTTGGTCTT 2478  
Db 2245 CTCTAAACATATTTGAATTTAGGGCCCTTAAGAACTTATGATCTTATTTGGTCTT 2304  
Qy 2479 TCTGAGGTTCAAGAGGGGTAGTAACTTCTCCAGGTCACACAGAGTCTGTGGTGGC 2538  
Db 2305 TCTGAGGTTCAAGAGGGGTAGTAACTTCTCCAGGTCACACAGAGTCTGTGGTGGC 2364  
Qy 2539 AGAAGCAAGCTAGCGCTGGGATTCAGTACATACACAGATGCTCTCTCTGATGCT 2598  
Db 2365 AGAAGCAAGCTAGCGCTGGGATTCAGTACATACACAGATGCTCTCTCTGATGCT 2424  
Qy 2599 TGGGCCCTGGGGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTCTCTCTGATGAG 2658  
Db 2425 TGGGCCCTGGGGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTCTCTCTGATGAG 2484  
Qy 2659 TCTGAGAGGGTCCCTGTGATTTGTGACACAGATGCTCATGATATACAGTACACACA 2718  
Db 2485 TCTGAGAGGGTCCCTGTGATTTGTGACACAGATGCTCATGATATACAGTACACACA 2544  
Qy 2719 GATGTACACACAGATGTACACATGCTCCAGGCCAGCTCTGATACCTGACCTGACCC 2778  
Db 2545 GATGTACACACAGATGTACACATGCTCCAGGCCAGCTCTGATACCTGACCTGACCC 2604  
Qy 2779 CAGCTTGGCCCCCTGCTGCTGTGTCTCAAGCAGAGTCCAACTCTCTCTGCTC 2838  
Db 2605 CAGCTTGGCCCCCTGCTGCTGTGTCTCAAGCAGAGTCCAACTCTCTCTGCTC 2664  
Qy 2839 CTTTCCCAACCACTGCTGAGCTTCTGAGCAGACAGGATACCTTGGCTGACCGGTGT 2898  
Db 2665 CTTTCCCAACCACTGCTGAGCTTCTGAGCAGACAGGATACCTTGGCTGACCGGTGT 2724  
Qy 2899 GTGACCGGCTCTCAACCAAGGACAGCCCGCCACATGATCTCGGTATACATATCAAT 2958  
Db 2725 GTGACCGGCTCTCAACCAAGGACAGCCCGCCACATGATCTCGGTATACATATCAAT 2784  
Qy 2959 AAAAGTGGTTTGTAAACAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 3006  
Db 2785 AAAAGTGGTTTGTAAACAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 2832

RESULT 5  
ABX72001  
ID ABX72001 standard; DNA; 2832 BP.  
XX ABX72001;  
AC ABX72001;  
XX 12-MAR-2003 (first entry)  
DT  
XX  
XX DNA encoding human tumour endothelial marker TEM 2.  
DE  
XX Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
KW

KM Tumour endothelial marker; normal endothelial marker; PEM;  
KM pan-endothelial marker; polycystic kidney disease; psoriasis;  
KM diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
KM neovascularization; immune response; cytoskeletal; gene;  
KM opthalmological; antineoplastic; antirheumatic; antipsoriatic; ds.  
OS Homo sapiens.  
PN WO20028374-A2.  
XX 24-OCT-2002.  
PD 10-APR-2002; 2002WO-US008253.  
PF 11-APR-2001; 2001US-0282850P.  
XX 06-FEB-2002; 2002US-0354262P.  
PR (UJJO) UNIV JOHNS HOPKINS.  
PA Carson-Walter E, St Croix B, Kinzler KM, Vogelstein B;  
PI WPI; 2003-093016/08.  
DR P-PSDB; ABUS4429.  
XX New purified human transmembrane protein, designated as tumor endothelial  
PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors, or  
PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
PT psoriasis.  
XX Disclosure; Page 113-114; 374pp; English.  
XX The present invention relates to a novel method for the isolation of  
CC endothelial cells (ECs), and the identification of genes expressed in  
CC normal and tumor ECs. Tumour endothelial marker (TEM), normal  
CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
CC identified in human ECs. The human EC marker proteins and the  
CC polynucleotide sequences encoding them are useful for detecting,  
CC diagnosing or treating tumors as well as polycystic kidney disease,  
CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
CC useful for inhibiting neovascularization or tumour angiogenesis, for  
CC inducing an immune response to tumour endothelial cells in a patient, or  
CC for identifying candidate drugs for treating tumors. The present  
CC sequence represents a human TEM or NEM gene of the invention  
XX  
SQ Sequence 2832 BP, 602 A; 891 C; 727 G; 612 T; 0 U; 0 Other;  
Query Match 89.5%; Score 2757; DB 7; Length 2832;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 199 CAGCCCGGAGCCATGATGAAGACTTGTTCAGCGGGAATGCAAGCTGATGTCGCCGC 258  
DB 25 CAGCCCGGAGCCATGATGAAGACTTGTTCAGCGGGAATGCAAGCTGATGTCGCCGC 84  
QY 259 AAAAATCATACCGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 318  
DB 85 AAAAATCATACCGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 144  
QY 319 TCTGCTTCTCAATGCGCGCTTTGAGAGCAAGTACACACCAACCATGAGAGTCTTCAC 378  
DB 145 TCTGCTTCTCAATGCGCGCTTTGAGAGCAAGTACACACCAACCATGAGAGTCTTCAC 204  
QY 379 CGTAAAGTATCAACATCGCGGCGAGATGTAACAGCTGACATCTTGATATCTTGGC 438  
DB 205 CGTAAAGTATCAACATCGCGGCGAGATGTAACAGCTGACATCTTGATATCTTGGC 264  
QY 439 AACCAACCCCTTCCCGCATGCGAGGCTGTCATCTCAACAGGAGATGTTCACTCTG 498  
DB 265 AACCAACCCCTTCCCGCATGCGAGGCTGTCATCTCAACAGGAGATGTTCACTCTG 324  
QY 499 GTGTTCAAGCTGATTAACCGGAGATCTTTCATGAGTCAAGCGCTTCAAGAGCATTC 558  
DB 325 GTGTTCAAGCTGATTAACCGGAGATCTTTCATGAGTCAAGCGCTTCAAGAGCATTC 384

QY 559 CTGAGGTCAGATCTCTGCTGAAGAACAGACCAAGAGCGGCGAGCTGCGCATGTC 618  
DB 385 CTGAGGTCAGATCTCTGCTGAAGAACAGACCAAGAGCGGCGAGCTGCGCATGTC 444  
QY 619 ATCTGTGGCAACAAGAACGACCAAGGAGCTGTGCGGAGTGCACCAACCGAGGC 678  
DB 445 ATCTGTGGCAACAAGAACGACCAAGGAGCTGTGCGGAGTGCACCAACCGAGGC 504  
QY 679 GAGCTGTGTGTGTGGGCGAGAGAACTCCGCTTACTTCCAGGTGTGGCCAAAGAAC 738  
DB 505 GAGCTGTGTGTGTGGGCGAGAGAACTCCGCTTACTTCCAGGTGTGGCCAAAGAAC 564  
QY 739 ACCAAGGAGCAAGATGTTCTAGTCTCTTCAAGTGTGCAAGTGTGCAAGAGATG 798  
DB 565 ACCAAGGAGCAAGATGTTCTAGTCTCTTCAAGTGTGCAAGTGTGCAAGAGATG 624  
QY 799 AGCCCGCCCTGATCCGCAAGATCTCCGTGCAAGTACGAGTGCAGCTTCCACCCAGGCCC 858  
DB 625 AGCCCGCCCTGATCCGCAAGATCTCCGTGCAAGTACGAGTGCAGCTTCCACCCAGGCCC 684  
QY 859 TTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGATGCTGTGCTTGGCCGC 918  
DB 685 TTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGATGCTGTGCTTGGCCGC 744  
QY 919 CGCCCGAGCGTCAACAGTGAACCTTCAAGTACCAAGGCGCAAGGCTTCCGGAAGGCGAG 978  
DB 745 CGCCCGAGCGTCAACAGTGAACCTTCAAGTACCAAGGCGCAAGGCTTCCGGAAGGCGAG 804  
QY 979 GCCCGTGAAGAGGCAAGTGAACCATCAGTGAAGCGAGTGTGGGCGGCGCTTGGC 1038  
DB 805 GCCCGTGAAGAGGCAAGTGAACCATCAGTGAAGCGAGTGTGGGCGGCGCTTGGC 864  
QY 1039 CAGTGCCTTCAAGGAGTGTGCGGCAATGCCACTGTGCGCATCTCCACCGAGGCCC 1098  
DB 865 CAGTGCCTTCAAGGAGTGTGCGGCAATGCCACTGTGCGCATCTCCACCGAGGCCC 924  
QY 1099 GGCAGAGCTCTTGTTCACAGCTTGAAGCAACCTGAGAGGCGGCGGCGGCGCTC 1158  
DB 925 GGCAGAGCTCTTGTTCACAGCTTGAAGCAACCTGAGAGGCGGCGGCGGCGCTC 984  
QY 1159 CGCAGATTCGCTGCTCTTCAACAGCTTCTGAGTCCGCTTGTGCAAGCTCTTGGTG 1218  
DB 985 CGCAGATTCGCTGCTCTTCAACAGCTTCTGAGTCCGCTTGTGCAAGCTCTTGGTG 1044  
QY 1219 GTTTCATCTCTCTGTGAGAGCAACATCTGCAAGCTTCAAGATTAGGCAAGACTCA 1278  
DB 1045 GTTTCATCTCTCTGTGAGAGCAACATCTGCAAGCTTCAAGATTAGGCAAGACTCA 1104  
QY 1279 AGTTACACCTTCTCTCTGAGAGCAACATCTGCAAGCTTCAAGATTAGGCAAGACTCA 1338  
DB 1105 AGTTACACCTTCTCTCTGAGAGCAACATCTGCAAGCTTCAAGATTAGGCAAGACTCA 1164  
QY 1339 TGCGTCAATGAGAGCTCTTGTGAGCAAGCTCAGAGTGAAGAGCAAGAGGCGAGAT 1398  
DB 1165 TGCGTCAATGAGAGCTCTTGTGAGCAAGCTCAGAGTGAAGAGCAAGAGGCGAGAT 1224  
QY 1399 GAGAAAGTCTCTCTCTCTGAGCAACATCTGCAAGCTTGTGAGGCTGAGGAGAG 1458  
DB 1225 GAGAAAGTCTCTCTCTCTGAGCAACATCTGCAAGCTTGTGAGGCTGAGGAGAG 1284  
QY 1459 ACTTCTCTCCAGGCTGCAACTTTCAGCTGTGCTTCAAGTGTGCAAGCTTCTCC 1518  
DB 1285 ACTTCTCTCCAGGCTGCAACTTTCAGCTGTGCTTCAAGTGTGCAAGCTTCTCC 1344  
QY 1519 ACCCGAGAGCAACAGTGTGCGGCGAGCTGCGCTGATGAGCCAGTGAAGTCT 1578  
DB 1345 ACCCGAGAGCAACAGTGTGCGGCGAGCTGCGCTGATGAGCCAGTGAAGTCT 1404  
QY 1579 GTGTTGAAGGAGGCGTGGCCACATCTTGAAGCAAGCGCCACCATTAAGCAAGCCCA 1638  
DB 1405 GTGTTGAAGGAGGCGTGGCCACATCTTGAAGCAAGCGCCACCATTAAGCAAGCCCA 1464

1639 CCTCGTGAACGGGTTCTGAGCTCTCTCTAGTCCCTCCGCGGAGAGTGTGCTTT 1698  
1465 CCTCTGACCGGCTCTCAGGCTCTCTCTAGGTCCTCGCGCGAGAGTGTGCTTT 1524  
1699 GTTGTGTGACAGCTGTTTCTGTATAGTATAGTAAATGGAATCAATTTGACTG 1758  
1525 GTTGTGTGACAGCTGTTTCTGTATAGTATAGTAAATGGAATCAATTTGACTG 1584  
1759 TAAAGACCTAGTACTCTCTCTGTGAGCGCCCTCAACCAAGTTCAGATCCAGCTCC 1818  
1585 TAAAGACCTAGTACTCTCTCTGTGAGCGCCCTCAACCAAGTTCAGATCCAGCTCC 1644  
1819 ACCCGGAGAGCTCTCTCTGTGAGCGCCCTCAACCAAGTTCAGATCCAGCTCC 1878  
1645 ACCCGGAGAGCTCTCTCTGTGAGCGCCCTCAACCAAGTTCAGATCCAGCTCC 1704  
1879 ACATTGACCTCCGCAATGAGCTCCAGGTTTACAGAAATTTGCAACAGCTGAGGTCG 1938  
1705 ACATTGACCTCCGCAATGAGCTCCAGGTTTACAGAAATTTGCAACAGCTGAGGTCG 1764  
1939 CAGGCCAGAGCTGCTTTTATGCTCCATTCAAGAGGATACACCGAGACTCG 1998  
1765 CAGGCCAGAGCTGCTTTTATGCTCCATTCAAGAGGATACACCGAGACTCG 1824  
1999 AGGGGACAGATGAGCAGCAGGCGCCACTTGTCTCCAGGAATTCAGGCTACAGCTC 2058  
1825 AGGGGACAGATGAGCAGCAGGCGCCACTTGTCTCCAGGAATTCAGGCTACAGCTC 1884  
2059 CACCTAGAACAGAGCTGCTCTTACTGTGCTGCTCTCTCAAGCAATTTATTAAGACCTAC 2118  
1885 CACCTAGAACAGAGCTGCTCTTACTGTGCTGCTCTCTCAAGCAATTTATTAAGACCTAC 1944  
2119 TGGGGTGTGGGTTTCACTGTGCTCTTGAAGAAACCAAGAGGTCCTGCTGCTGCTG 2178  
1945 TGGGGTGTGGGTTTCACTGTGCTCTTGAAGAAACCAAGAGGTCCTGCTGCTGCTG 2004  
2179 CGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2238  
2005 CGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2064  
2239 TGGGACCCAGAACTGAGCTGCGAGGAGATCCGACAGAAAAGCTCAGGGGCTCTTCTC 2298  
2065 TGGGACCCAGAACTGAGCTGCGAGGAGATCCGACAGAAAAGCTCAGGGGCTCTTCTC 2124  
2299 TTTGCTCCGGGATTTGGGCTATGCTGAGTACCAACATGATCTCAGGCTATGAGGTTTGA 2358  
2125 TTTGCTCCGGGATTTGGGCTATGCTGAGTACCAACATGATCTCAGGCTATGAGGTTTGA 2184  
2359 ACCCATTAACCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2418  
2185 ACCCATTAACCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2244  
2419 CTCTAAACATTTGAAATTTAGGCTCCTTAAGAACTTTAGTATCTTATTTGGGCTT 2478  
2245 CTCTAAACATTTGAAATTTAGGCTCCTTAAGAACTTTAGTATCTTATTTGGGCTT 2304  
2479 TCTGAGTTTCAGAGGGGTAAGTAACTTCTCAGAGTACACAGAAAGCTGTGGGTGGC 2538  
2305 TCTGAGTTTCAGAGGGGTAAGTAACTTCTCAGAGTACACAGAAAGCTGTGGGTGGC 2364  
2539 AGAAGCAAGCTAGAGCTGCGGCTTTCAGTACATACACAGATGCTCCTCTCTTGTATGCT 2598  
2365 AGAAGCAAGCTAGAGCTGCGGCTTTCAGTACATACACAGATGCTCCTCTCTTGTATGCT 2424  
2599 TGGGCTCTGGGGCTTTCAGGGCTTTGGGACATCTTGTCTCAACCTCTCTCTATATAG 2658  
2425 TGGGCTCTGGGGCTTTCAGGGCTTTGGGACATCTTGTCTCAACCTCTCTCTATATAG 2484  
2659 TCTGTAGAGGTCCTCTGTAGATTTGTGTACACATGAGCCATGATATATCAAGTACACA 2718  
2485 TCTGTAGAGGTCCTCTGTAGATTTGTGTACACATGAGCCATGATATATCAAGTACACA 2544  
2719 GATGTACACACAGATGTACACATGCTCCAGGCCCAAGCTCTGTGATACCTGCACTGCAACC 2778

2545 GATGTACACACAGATGTACACATGCTCCAGGCCCAAGCTCTGTGATACCTGCACTGCAACC 2604  
2779 CAGCTTTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2838  
2605 CAGCTTTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2664  
2839 CTTTCCCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2898  
2665 CTTTCCCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2724  
2899 GTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2958  
2725 GTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2784  
2959 AAAAGTGGGTTTGTATCAAAAAA 3006  
2785 AAAAGTGGGTTTGTATCAAAAAA 2832

RESULT 6  
ABK71563  
ID ABK71563 standard; cDNA; 3427 BP.  
XX  
AC ABK71563;  
XX  
DT 30-JUL-2002 (first entry)  
XX  
DE Human dithp polynucleotide #29.  
XX  
KW Human; dithp; diagnostic and therapeutic polynucleotide; gene; ss; bone;  
cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;  
inflammatory disorder; viral infection; bacterial infection; seizure;  
fungal infection; parasitic infections; developmental disorder; breast;  
endocrine disorder; metabolic disorder; neurological disorder; cervix;  
gastrointestinal disorder; transport disorder; gene therapy; kidney;  
adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;  
skin; testis; thymus.  
XX  
OS Homo sapiens.  
XX  
PN WO200220754-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 29-AUG-2001; 2001WO-US027127.  
XX  
PR 05-SEP-2000; 2000US-0229747P.  
PR 05-SEP-2000; 2000US-0229748P.  
PR 05-SEP-2000; 2000US-0229749P.  
PR 05-SEP-2000; 2000US-0229750P.  
PR 05-SEP-2000; 2000US-0229751P.  
PR 05-SEP-2000; 2000US-023053P.  
PR 06-SEP-2000; 2000US-0230505P.  
PR 06-SEP-2000; 2000US-0230514P.  
PR 06-SEP-2000; 2000US-0230515P.  
PR 06-SEP-2000; 2000US-0230517P.  
PR 06-SEP-2000; 2000US-0230518P.  
PR 06-SEP-2000; 2000US-0230519P.  
PR 06-SEP-2000; 2000US-0230595P.  
PR 06-SEP-2000; 2000US-0230597P.  
PR 06-SEP-2000; 2000US-0230598P.  
PR 06-SEP-2000; 2000US-0230599P.  
PR 06-SEP-2000; 2000US-0230610P.  
PR 06-SEP-2000; 2000US-023065P.  
PR 06-SEP-2000; 2000US-0230988P.  
PR 07-SEP-2000; 2000US-0230951P.  
PR 07-SEP-2000; 2000US-0231163P.  
PR 07-SEP-2000; 2000US-0231167P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JT,

PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;  
PI Moriama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;  
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Dalfio A;  
PI Marwana R, Chen AU, Chang SC, Au AP, Imman RR;

DR WPI: 2002-383054/41.  
DR P-PSDB: ABG59971.

PT An isolated polynucleotide useful in diagnostics and therapeutics.

PS Claim 1: Page 419-420; 686pp; English.

XX The invention relates to human diagnostic and therapeutic (dithp)  
XX polynucleotides and their associated polypeptides (dithp polypeptides).  
CC The sequences of the invention are used in the treatment and diagnosis of  
CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers  
CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,  
CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or  
CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,  
CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal  
CC infections, parasitic infections, developmental disorders (e.g. anaemia,  
CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),  
CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders  
CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,  
CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal  
CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders  
CC (e.g. myotonic dystrophy, catatonla, peripheral neuropathy). Sequences  
CC ABK7135-ABK71809 represent human dithp polynucleotides of the invention  
XX

Sequence 3427 BP; 682 A; 1145 C; 935 G; 665 T; 0 U; 0 Other;

Query Match 83.4%; Score 2571; DB 6; Length 3427;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2961; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy	26	GCCGCGCCGAGCCGCGGCTCCGAGCAGCGAGGAGGATCCCGGCGAGTGAACCGGG	85
Db	459	GCCGCGCCGAGCCGCGGCTCCGAGCAGCGAGGAGGATCCCGGCGAGTGAACCGGG	518
Qy	86	AGCCACCAAGAGCTCTGGAGAGCTCGGCGGCTGAGAGCAGAGCACTCCCGGAGCTCC	145
Db	519	AGCCACCAAGAGCTCTGGAGAGCTCGGCGGCTGAGAGCAGAGCACTCCCGGAGCTCC	578
Qy	146	CGGCGCTTCCAGGAGCTCTTGAAGCCTGCAAGGCGCGCGCCGATTTCCAGCCGCC	205
Db	579	CGGCGCTTCCAGGAGCTCTTGAAGCCTGCAAGGCGCGCGCCGATTTCCAGCCGCC	638
Qy	206	GAGCGATGATGAGACTTGTCCAGCGGGAACCTGAGCGCTGAGTGGCGGCAAAAAC	265
Db	639	GAGCGATGATGAGACTTGTTCAGCGGGAACCTGAGCGCTGAGTGGCGGCAAAAAC	698
Qy	266	CATACCGAGTGTGTGTGAGGCTCTCGGGTGGGAGAGCTCATCTGTGTGCT	325
Db	699	CATACCGAGTGTGTGTGAGGCTCTCGGGTGGGAGAGCTCATCTGTGTGCT	758
Qy	326	TCCTCATGCGCGCTTTGAGAGCAGTACACCAACCATGAGGACTTCCACGGTAAG	385
Db	759	TCCTCATGCGCGCTTTGAGAGCAGTACACCAACCATGAGGACTTCCACGGTAAG	818
Qy	386	TATACACATCGCGGCGAGACATGTAACGACTCGACATCTGGATTCCTTGGCAACAC	445
Db	819	TATACACATCGCGGCGAGACATGTAACGACTCGACATCTGGATTCCTTGGCAACAC	878
Qy	446	CTTTCCTCCCGCAGTCGAGGCTGTTCATCTCAAGGGAGTGTCTTCATCTGTGTTC	505
Db	879	CTTTCCTCCCGCAGTCGAGGCTGTTCATCTCAAGGGAGTGTCTTCATCTGTGTTC	938
Qy	506	GCTTGATTAACCGGAGTCTTTCATGAGTCAAGCGCTTCAAGAGAGATCTGTGAGG	565
Db	939	GCTTGATTAACCGGAGTCTTTCATGAGTCAAGCGCTTCAAGAGAGATCTGTGAGG	998
Qy	566	TCAAGTCTGCTGAGAAAGAACAGACCAAGAGCGCGGAGCTGCCATGTGATCTGTG	625

Db	999	TCAAGTCTGCTGAGAAAGAACAGACCAAGAGCGCGGAGCTGCCATGTGATCTGTG	1058
Qy	626	GCAACAGAAAGACACACAGGAGGTGTGCGCCAGGTGCCACACCGAGGCCAGTGTG	685
Db	1059	GCAACAGAAAGACACACAGGAGGTGTGCGCCAGGTGCCACACCGAGGCCAGTGTG	1118
Qy	686	TGATGTGCGGCGAGAGAACTCCGCTACTTTCGAGGTGTGCGGCAAGAAACCAACG	745
Db	1119	TGATGTGCGGCGAGAGAACTCCGCTACTTTCGAGGTGTGCGGCAAGAAACCAACG	1178
Qy	746	TGACAGAGATGTTTCTAGTGTCTTTCAGATGAGCAAGCTGCCACAGAGATGAGCCCG	805
Db	1179	TGACAGAGATGTTTCTAGTGTCTTTCAGATGAGCAAGCTGCCACAGAGATGAGCCCG	1238
Qy	806	CCCTGATTCGAGATCTCCGTGAGTACGCTGACGCTTTCACCCAGGCGCTTTCGCA	865
Db	1239	CCCTGATTCGAGATCTCCGTGAGTACGCTGACGCTTTCACCCAGGCGCTTTCGCA	1298
Qy	866	TGCGCGGCTCAAGAGATGAGAGCGCTATGAGCATGTGCTGCGCTTGCAGCGGCCCA	925
Db	1299	TGCGCGGCTCAAGAGATGAGAGCGCTATGAGCATGTGCTGCGCTTGCAGCGGCCCA	1358
Qy	926	GCGTCAACAGTACCTCAAGTACATCAAGGCAAGTCTTTCGAGAAAGCCAGGCGCGTG	985
Db	1359	GCGTCAACAGTACCTCAAGTACATCAAGGCAAGTCTTTCGAGAAAGCCAGGCGCGTG	1418
Qy	986	AGAGGACAAAGTGCACATTCAGTACGAGAGGATGCTGCGGCGCGGCTTGGCCAGTGC	1045
Db	1419	AGAGGACAAAGTGCACATTCAGTACGAGAGGATGCTGCGGCGCGGCTTGGCCAGTGC	1478
Qy	1046	TTTCAGGAGAGTGGCCCAAGATGCGCATGCGCATCTCCCAAGAGGCCCGGAGCA	1105
Db	1479	TTTCAGGAGAGTGGCCCAAGATGCGCATGCGCATCTCCCAAGAGGCCCGGAGCA	1538
Qy	1106	GTCCTTGTTCACAGACTTTAGGACACAGCTGAGAGGCCCGGCGGCTTGCCTCCGACAT	1165
Db	1539	GTCCTTGTTCACAGACTTTAGGACACAGCTGAGAGGCCCGGCGGCTTGCCTCCGACAT	1598
Qy	1166	TGCTGTGCTTTCACAGCTTTTCTGAGTCCGCTTTCACAGCTCTTGTGTGTTCAT	1225
Db	1599	TGCTGTGCTTTCACAGCTTTTCTGAGTCCGCTTTCACAGCTCTTGTGTGTTCAT	1658
Qy	1226	CTCCTCTGTGGAGAGACATCTCTGAGGCTCAAGAGTTAAGGAGAGCTCAAGTTACA	1285
Db	1659	CTCCTCTGTGGAGAGACATCTCTGAGGCTCAAGAGTTAAGGAGAGCTCAAGTTACA	1718
Qy	1286	CCTTCTCTCTCGGGTTGGAAGAAATGTTGATGCAAGAGGAGTGAAGATTGCTGCTCA	1345
Db	1719	CCTTCTCTCTCGGGTTGGAAGAAATGTTGATGCAAGAGGAGTGAAGATTGCTGCTCA	1778
Qy	1346	TATGAGGCTTCTGGGACAAAGCTCAGAGTGAAGAAACACAGAAAGCAGATGAAG	1405
Db	1779	TATGAGGCTTCTGGGACAAAGCTCAGAGTGAAGAAACACAGAAAGCAGATGAAG	1838
Qy	1406	GTCCTCTCTCTCGGGATGAACACCCAGCTTGTGTTGGTGGGCGAGTGGGAACTTCTC	1465
Db	1839	GTCCTCTCTCTCGGGATGAACACCCAGCTTGTGTTGGTGGGCGAGTGGGAACTTCTC	1898
Qy	1466	TCCAGGCTTCGACACTTTCAGCTCTGATGAGCTGTGACCCCTTCCACCCCA	1525
Db	1899	TCCAGGCTTCGACACTTTCAGCTCTGATGAGCTGTGACCCCTTCCACCCCA	1958
Qy	1526	GCAACACACAAAGTGGCCCCAGCTGCGCTGACATTGAGCCAGTGAAGCTGTGTGTG	1585
Db	1959	GCAACACACAAAGTGGCCCCAGCTGCGCTGACATTGAGCCAGTGAAGCTGTGTGTG	2018
Qy	1586	AAAGGGGCGTGGCCACACCTCTTAAGACACGCGCCACACTTAAGACACGCGCCACTC	1645
Db	2019	AAAGGGGCGTGGCCACACCTCTTAAGACACGCGCCACACTTAAGACACGCGCCACTC	2078
Qy	1646	ACCGGTTCTCAGGCTCTCTCTTGAAGTCTCTCGCGCCGACAGTGTGCTTGTGTG	1705
Db	2079	ACCGGTTCTCAGGCTCTCTCTTGAAGTCTCTCGCGCCGACAGTGTGCTTGTGTG	2138



QY	1706	TTGACAGCTGTTTCCGTCACATGTAAGTAAGTAAGAAATGGAATCACTTGTACTGTAAAAAGC	1765
Db	2139	TTGACAGCTGTTTCCGTCACATGTAAGTAAGTAAGAAATGGAATCACTTGTACTGTAAAAAGC	2198
QY	1766	CTAGTGACTCCCTCTTGGCCAGGCGCTCACCCAGTTCAAGATCCACGGCGCTCCACCCGGG	1825
Db	2199	CTAGTGACTCCCTCTTGGCCAGGCGCTCACCCAGTTCAAGATCCACGGCGCTCCACCCGGG	2258
QY	1826	AGCCTTCTCTCTCTGCTCCCAAACAGGGTTTCCGTGGCTGTTTGCAGCTAGACATTGA	1885
Db	2259	AGCCTTCTCTCTCTGCTCCCAAACAGGGTTTCCGTGGCTGTTTGCAGCTAGACATTGA	2318
QY	1886	CTCCGCAATTAGCTCCACGGTTTACAGACAAATTGCACAAGCGTGGGGTGGAGGCGCA	1945
Db	2319	CTCCGCAATTAGCTCCACGGTTTACAGACAAATTGCACAAGCGTGGGGTGGAGGCGCA	2378
QY	1946	GGACTGCTTTTTTTTAAATGCTCCATTCCACAGAGATACCAACCGACCTCGAGAGGGAC	2005
Db	2379	GGACTGCTTTTTTTTAAATGCTCCATTCCACAGAGATACCAACCGACTCGAGAGGGAC	2438
QY	2006	ACGATGAGACACCAAGGCCCCACCTTGTGCCCTAGAAATTGACGGGTAAAGTCCACCTAG	2065
Db	2439	ACGATGAGACACCAAGGCCCCACCTTGTGCCCTAGAAATTGACGGGTAAAGTCCACCTAG	2498
QY	2066	AAACGAGCTGCCCTTACTGTGTGCTGCTTCAAGCATTTATTAAGCACTTACGTGGTGC	2125
Db	2499	AAACGAGCTGCCCTTACTGTGTGCTGCTTCAAGCATTTATTAAGCACTTACGTGGTGC	2558
QY	2126	TGGGTTCACCTGTGTCTTAGAAACCAAGAGGTCCTCCAGTCTGTGGCTCTGCGCCCT	2185
Db	2559	TGGGTTCACCTGTGTCTTAGAAACCAAGAGGTCCTCCAGTCTGTGGCTCTGCGCCCT	2618
QY	2186	GCTGCCCAACCACTTCTGCACACACACGCGTGGGAGCGCGGAGAGACGTGGGAC	2245
Db	2619	GCTGCCCAACCACTTCTGCACACACACGCGTGGGAGCGCGGAGAGACGTGGGAC	2678
QY	2246	CAGAACTGAAGCTTGGAGAGGATCCGACAGAAAAGCTCAGGGCGGGTCTTCT - CCTGTGC	2304
Db	2679	CAGAACTGAAGCTTGGAGAGGATCCGACAGAAAAGCTCAGGGCGGGTCTTCT - CCTGTGC	2738
QY	2305	CCGGATTGGGGTATGCTGGGGATCCACATGTACTCAGGCATGGTGGGTTTGAACCAT	2364
Db	2739	CCGGATTGGGGTATGCTGGGGATCCACATGTACTCAGGCATGGTGGGTTTGAACCAT	2798
QY	2365	AAACCAAAAGCCCTTGTCATCAGCTCTTAAACAAGTATATTTGTATTTTAACTCTCTAA	2424
Db	2799	AAACCAAAAGCCCTTGTCATCAGCTCTTAAACAAGTATATTTGTATTTTAACTCTCTAA	2858
QY	2425	ACATTTGAAGTTTAAAGGCCCTTAAGGAACCTTAGTATCTTCTAATTGGGCTTCTGAG	2484
Db	2859	ACATTTGAAGTTTAAAGGCCCTTAAGGAACCTTAGTATCTTCTAATTGGGCTTCTGAG	2918
QY	2485	GTTCCAGAGAGGTAAGTAACTTCTCCACAG - TCAACACGACAGTCTGAGGTGCAGAAAG	2543
Db	2919	GTTCCAGAGAGGTAAGTAACTTCTCCACAGTTCACACGACAGTCTGAGGTGCAGAAAG	2978
QY	2544	CAAGTGAAGCTGGGCACTTCACTACATACCAAGATGTGCTCCCTCTCTTGAATGCTTGGCC	2603
Db	2979	CAAGTGAAGCTGGGCACTTCACTACATACCAAGATGTGCTCCCTCTCTTGAATGCTTGGCC	3038
QY	2604	CCTGGGGGCTTCAAGGGCTTTGGGACATTTGTCTCAACCCCTCTCCCTAGATCAGTCTGT	2663
Db	3039	CCTGGGGGCTTCAAGGGCTTTGGGACATTTGTCTCAACCCCTCTCCCTAGATCAGTCTGT	3098
QY	2664	GAGGGTCCCTGTAGATATTGTGTACACCATGCCCCATGTATTAACAAGTACACACAGATGT	2723
Db	3099	GAGGGTCCCTGTAGATATTGTGTACACCATGCCCCATGTATTAACAAGTACACACAGATGT	3158
QY	2724	ACACACAGATTTACACATGCTCCAGGCCCAAGCTCTGCATACCTGCACTGCAACCCAGCC	2783
Db	3159	ACACACAGATTTACACATGCTCCAGGCCCAAGCTCTGCATACCTGCACTGCAACCCAGCC	3218

QY	2784	TTGGCCCTGCTGGTGTGTGCTCAAGCAGCAGCTCAACCCGTGCTGTGCCCTTC	2843
Db	3219	TTGGCCCTGCTGGTGTGTGCTCAAGCAGCAGCTCAACCCGTGCTGTGCCCTTC	3278
QY	2844	CCCAACCACTGCTGAGGCTTTCTGAGCAGACCAAGTACCTTGGCTGCACCGGTGTGTGC	2903
Db	3279	CCCAACCACTGCTGAGGCTTTCTGAGCAGACCAAGTACCTTGGCTGCACCGGTGTGTGC	3338
QY	2964	CCGCTTCACCCAGGACAGCCCGCCACCACTGATCTCCGTATCACTATCAATTAAG	2966
Db	3339	CCGCTTCACCCAGGACAGCCCGCCACCACTGATCTCCGTATCACTATCAATTAAG	3398
QY	2964	TTGGTTTGTTCACAAAAA	2989
Db	3399	TTGGTTTGTTCACAAAAA	3424
RESULT 7			
ABSS4740			
ID	ABSS4740	standard; DNA; 11221 BP.	
XX	ABSS4740;		
AC	04-DEC-2002	(first entry)	
XX			
DE	Genomic DNA encoding human Ras-like protein.		
XX			
KW	Human; Ras-like; cancer; apoptosis; cell proliferation; AIDS; allergy;		
KW	acquired immunodeficiency syndrome; Alzheimer's disease; osteoporosis;		
KW	Parkinson's disease; aplastic anemia; atherosclerosis; inflammation;		
KW	bacterial infection; viral infection; fungal infection; gene therapy;		
KW	parasitic infection; protozoal infection; helminthic infection; SNP;		
KW	Ras-inhibitor; Ras-stimulator; single nucleotide polymorphism; gene; ds.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	location/Qualifiers	
FT	CDS	3000..8223	
FT		/*tag= a	
FT		/*product= "Human Ras-like protein"	
FT	exon	3000..3270	
FT		/*tag= b	
FT		/number=1	
FT	intron	3271..7693	
FT		/*tag= c	
FT		/number=1	
FT	variation	replace(3951,T)	
FT		/*tag= d	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	
FT		replace(4127,T)	
FT		/*tag= e	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	
FT		replace(4157,A)	
FT		/*tag= f	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	
FT		replace(4513,T)	
FT		/*tag= g	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	
FT		replace(6894,C)	
FT		/*tag= h	
FT	exon	7694..8220	
FT		/*tag= i	
FT		/number=2	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	
FT		replace(8409,C)	
FT		/*tag= j	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	
FT		replace(8437,G)	
FT		/*tag= k	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	
FT		replace(8579,T)	
FT		/*tag= l	
FT	variation	/*standard_name= "Single nucleotide polymorphism (SNP) "	



FT variation replace(10292,G)  
 FT /\*tag= m  
 FT /standard\_name= "Single nucleotide polymorphism (SNP)"  
 FT variation replace(10792,G)  
 FT /\*tag= n  
 FT /standard\_name= "Single nucleotide polymorphism (SNP)"  
 PN MO20262849-A2.  
 XX 15-AUG-2002.  
 PD 28-JAN-2002; 2002MO-US002301.  
 PF 08-FEB-2001; 2001US-00778963.  
 XX (PEKE ) PE CORP NY.  
 PA Neelam B, Ketchum KA, Di Francesco V, Beasley EM;  
 PI WPI: 2002-706901/76.  
 DR P-PSDB: ABG70951.  
 XX New human Ras-like polypeptides, useful for treating diseases associated  
 PT with an increase in apoptosis or cell proliferation e.g., cancer.  
 PS Claim 4; Fig 3; 73pp: English.  
 CC The present invention relates to a new polypeptide comprising a fully  
 CC defined 266-amino acid sequence or its fragment, orthologue or allelic  
 CC variant. The allelic variant or orthologue is encoded by a nucleic acid  
 CC molecule that hybridises under stringent conditions to the opposite  
 CC strand of a nucleic acid molecule with a fully defined sequence of 3082  
 CC or 11221 bp. The pharmaceutical composition comprising the agent that  
 CC binds to the polypeptide is useful for treating a disease or condition  
 CC mediated by a human Ras-like protein e.g. cancer. Likewise, the  
 CC polypeptide is useful for identifying modulators of Ras-like protein  
 CC activity or expression by screening a compound for the ability to  
 CC stimulate or inhibit interaction between the Ras-like protein and an  
 CC interacting molecule or Ras-like activity. The polypeptide is useful for  
 CC diagnosing a disease, or predisposition to a disease, or treating  
 CC diseases associated with an increase in apoptosis or cell proliferation  
 CC e.g. AIDS (acquired immunodeficiency syndrome), Alzheimer's disease,  
 CC Parkinson's disease, aplastic anaemia, allergies, atherosclerosis,  
 CC osteoporosis or bacterial, viral, fungal, parasitic, protozoal or  
 CC helminthic infections, and inflammation. The present nucleic acid  
 CC sequence encodes the human Ras-like protein of the invention  
 CC  
 SQ Sequence 11221 BP; 2512 A; 3108 C; 3028 G; 2573 T; 0 U; 0 Other;  
 Query Match 79.5%; Score 2449; DB 6; Length 11221;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 779 CCAAGCTGCCACAGAGATGAGCCCGCCCTGTCATCCGCAAGATCTCCGTCAGTACGGTG 838  
 DB CCAAGCTGCCACAGAGATGAGCCCGCCCTGTCATCCGCAAGATCTCCGTCAGTACGGTG 8050  
 QY 7991 CCAAGCTGCCACAGAGATGAGCCCGCCCTGTCATCCGCAAGATCTCCGTCAGTACGGTG 8050  
 DB 839 ACGCCTTCCACCCGAGGCCCTTCTGATGCGCGCGGTCAAGAGATGAGCGCTATGCGCA 898  
 DB ACGCCTTCCACCCGAGGCCCTTCTGATGCGCGCGGTCAAGAGATGAGCGCTATGCGCA 8110  
 QY 899 TGTGCTGGCCCTTGGCCCGCCCGCCGAGCTGCAACATGACCTTCAATATCAAGGCCA 958  
 DB TGTGCTGGCCCTTGGCCCGCCCGCCGAGCTGCAACATGACCTTCAATATCAAGGCCA 8170  
 QY 8111 TGTGCTGGCCCTTGGCCCGCCCGCCGAGCTGCAACATGACCTTCAATATCAAGGCCA 8170  
 DB 959 AGTCTCTTGGGAGGAGCGCCGCTGAGAGGAGCAAGTGCACCATCAAGTACGAGCGAGG 1018  
 DB AGTCTCTTGGGAGGAGCGCCGCTGAGAGGAGCAAGTGCACCATCAAGTACGAGCGAGG 8230  
 QY 8171 AGTCTCTTGGGAGGAGCGCCGCTGAGAGGAGCAAGTGCACCATCAAGTACGAGCGAGG 8230  
 DB 1019 ATGCTGGGCGCGGCGCTTGGCCGAGTGCCTTCAAGGAGGTGGCCCGAGATGCCACTGTGCG 1078  
 DB ATGCTGGGCGCGGCGCTTGGCCGAGTGCCTTCAAGGAGGTGGCCCGAGATGCCACTGTGCG 8230  
 QY 1079 CATCTCCCGACCGAGGCCCGCGGACGAGCTCTTGTTCACACACCTTAGGACACCACTGGA 1138  
 DB CATCTCCCGACCGAGGCCCGCGGACGAGCTCTTGTTCACACACCTTAGGACACCACTGGA 8350  
 QY 8291 CATCTCCCGACCGAGGCCCGCGGACGAGCTCTTGTTCACACACCTTAGGACACCACTGGA 8350  
 DB 1139 GAGCCCGCGGCGCTGGGCGCTCGGACATTCGCTGCTTCAAGCTTCTTCTGAGTCCGC 1198  
 DB GAGCCCGCGGCGCTGGGCGCTCGGACATTCGCTGCTTCTTCAAGCTTCTTCTGAGTCCGC 8410  
 QY 1199 TTGTCCACAGCTCTTGGGAGTTCATCTCTCTGTGAGGAGACACATCTGACGCTC 1258  
 DB TTGTCCACAGCTCTTGGGAGTTCATCTCTCTGTGAGGAGACACATCTGACGCTC 8470  
 QY 8411 TTGTCCACAGCTCTTGGGAGTTCATCTCTCTGTGAGGAGACACATCTGACGCTC 8470  
 DB 1259 AAGGTTAGGAGAGACTCAAGTTACACCTTCTCTCTGAGGAGTGGAGAAATGTTGAT 1318  
 DB AAGGTTAGGAGAGACTCAAGTTACACCTTCTCTCTGAGGAGTGGAGAAATGTTGAT 8530  
 QY 1319 GCCAGAGGAGGAGATTCGCTGCTCATATGAGCTCTCTGAGGAGCAAGCTCAGAGTAA 1378  
 DB GCCAGAGGAGGAGATTCGCTGCTCATATGAGCTCTCTGAGGAGCAAGCTCAGAGTAA 8590  
 QY 8591 GCCAGAGGAGGAGATTCGCTGCTCATATGAGCTCTCTGAGGAGCAAGCTCAGAGTAA 8590  
 DB 1379 AAGGACACAGAGGCGAGATGAGAAAGTCTCTCTCTGAGGAGTAAACCCAGCTTGG 1438  
 DB AAGGACACAGAGGCGAGATGAGAAAGTCTCTCTCTGAGGAGTAAACCCAGCTTGG 8650  
 QY 8591 AAGGACACAGAGGCGAGATGAGAAAGTCTCTCTCTGAGGAGTAAACCCAGCTTGG 8650  
 DB 1439 TTGGGTTGGCAGCTGGGAGAACTTCTCTCTGAGGAGTAAACCCAGCTTGG 1498  
 DB TTGGGTTGGCAGCTGGGAGAACTTCTCTCTGAGGAGTAAACCCAGCTTGG 8710  
 QY 8651 TTGGGTTGGCAGCTGGGAGAACTTCTCTCTGAGGAGTAAACCCAGCTTGG 8710  
 DB 1499 CTGCTCTGCAACCCCTCCACCCCGGACACACACAAGTTGGCCCGAGCTGGCGCTG 1558  
 DB CTGCTCTGCAACCCCTCCACCCCGGACACACACAAGTTGGCCCGAGCTGGCGCTG 8770  
 QY 8771 CTGCTCTGCAACCCCTCCACCCCGGACACACACAAGTTGGCCCGAGCTGGCGCTG 8770  
 DB 1559 ACATTGAGCCAGTGAATCTGTGTCTGTAAGGAGGCGGTGACCACTCTAGACACGCC 1618  
 DB ACATTGAGCCAGTGAATCTGTGTCTGTAAGGAGGCGGTGACCACTCTAGACACGCC 8830  
 QY 8771 ACATTGAGCCAGTGAATCTGTGTCTGTAAGGAGGCGGTGACCACTCTAGACACGCC 8830  
 DB 1619 CACCACTTAGACACAGCCCACTCTCTGACGCTCTCTCTAGTCTCTCTAGTCTCT 1678  
 DB CACCACTTAGACACAGCCCACTCTCTGACGCTCTCTCTAGTCTCTCTAGTCTCTCT 8890  
 QY 8831 CACCACTTAGACACAGCCCACTCTCTGACGCTCTCTCTAGTCTCTCTAGTCTCTCT 8890  
 DB 1679 CCGCCCGACAGTTGAGCTTGTGTGAGTGAAGCTTGTGTGAGTGAAGCTTGTGTGAGTGA 1738  
 DB CCGCCCGACAGTTGAGCTTGTGTGAGTGAAGCTTGTGTGAGTGAAGCTTGTGTGAGTGA 8950  
 QY 8891 CCGCCCGACAGTTGAGCTTGTGTGAGTGAAGCTTGTGTGAGTGAAGCTTGTGTGAGTGA 8950  
 DB 1739 AAATGAAATATCTGTAATGTAAGCTTGTGTGAGTGAAGCTTGTGTGAGTGAAGCTTGT 1798  
 DB AAATGAAATATCTGTAATGTAAGCTTGTGTGAGTGAAGCTTGTGTGAGTGAAGCTTGT 9010  
 QY 8951 AAATGAAATATCTGTAATGTAAGCTTGTGTGAGTGAAGCTTGTGTGAGTGAAGCTTGT 9010  
 DB 1799 AGTTCAATCCAGGCGCTTCCACCGGAGAGCGCTTCTCTGCTCTCCCAACAGAGTTTC 1858  
 DB AGTTCAATCCAGGCGCTTCCACCGGAGAGCGCTTCTCTGCTCTCCCAACAGAGTTTC 9070  
 QY 9011 AGTTCAATCCAGGCGCTTCCACCGGAGAGCGCTTCTCTGCTCTCCCAACAGAGTTTC 9070  
 DB 1859 CGTGCGCTGTTTGAGGCTAGACATTGACCTCCGCAATGAGCTCCAGGTTTACAGACAA 1918

```

Db      9071  |||||
          CGTGGCCTGTTTGACGTAGACATTGACCTCCGCACTTGAGCTCCACGGTTTACAGACA 9130
Qy      1919  TTGCAAGCGGTGGGTGGGAGGCGCAGAGCTGTTTTTTTATGTCCTCCATTTACAG 1978
Db      9131  TTGCAAGCGGTGGGTGGGAGGCGCAGAGCTGTTTTTTTATGTCCTCCATTTACAG 9190
Qy      1979  AGGATACACCGAGACTGGAGGGGACAGATGACACAGGGCCCACTTGTCCCTA 2038
Db      9191  AGGATACACCGAGACTGGAGGGGACAGATGACACAGGGCCCACTTGTCCCTA 9250
Qy      2039  GCAATTCAGGGTACAGCTCCACCTAGAACAGAGCTGCTCTACTGTCTCTCA 2098
Db      9251  GCAATTCAGGGTACAGCTCCACCTAGAACAGAGCTGCTCTACTGTCTCTCA 9310
Qy      2099  AGCATTATTAAGACCTACTGAGGTGCTGAGTTCACTGTCTAGAAACCAAGAGGT 2158
Db      9311  AGCATTATTAAGACCTACTGAGGTGCTGAGTTCACTGTCTAGAAACCAAGAGGT 9370
Qy      2159  CCCCAGTCTGGCTCTGCTGCCCCCTGCTGCTGCCACCACTTTCGACACAGCGGTG 2218
Db      9371  CCCCAGTCTGGCTCTGCTGCCCCCTGCTGCTGCCACCACTTTCGACACAGCGGTG 9430
Qy      2219  GGGAGCGGGGAGAGAGAGCTGGAGCCGAGAACTGAGCTGGAGGAGATCCGACAGAAA 2278
Db      9431  GGGAGCGGGGAGAGAGAGCTGGAGCCGAGAACTGAGCTGGAGGAGATCCGACAGAAA 9490
Qy      2279  GCTCAGGGGGGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338
Db      9491  GCTCAGGGGGGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9550
Qy      2339  TCAGGCAATGTTGGTTTGAACCCATTAACCAAGGCTTGTCTGCTGCTTAAACAG 2398
Db      9551  TCAGGCAATGTTGGTTTGAACCCATTAACCAAGGCTTGTCTGCTGCTTAAACAG 9610
Qy      2399  TATATTTGATTTTAACTCTCTTAAACATATTGAAGTTTGAAGGCTTAAAGAACTTA 2458
Db      9611  TATATTTGATTTTAACTCTCTTAAACATATTGAAGTTTGAAGGCTTAAAGAACTTA 9670
Qy      2459  GTGATCTCTATGAGGTCTTTCTGAGGTTCAAGAGGAGTAACTTCTCCAGGTAC 2518
Db      9671  GTGATCTCTATGAGGTCTTTCTGAGGTTCAAGAGGAGTAACTTCTCCAGGTAC 9730
Qy      2519  ACAGCAAGTCTGTTGGTGGCAGAGCAAGCACTAGCGCTGGGCACTAGTACACACAT 2578
Db      9731  ACAGCAAGTCTGTTGGTGGCAGAGCAAGCACTAGCGCTGGGCACTAGTACACACAT 9790
Qy      2579  GTGCTCCCTCTCTTATGCTGAGGCTTGGGCTTCAAGGCTTGGGCACTTGTCT 2638
Db      9791  GTGCTCCCTCTCTTATGCTGAGGCTTGGGCTTCAAGGCTTGGGCACTTGTCT 9850
Qy      2639  CAACCTCTCTCTGATGCTGAGGCTTCTGTAGATATGTTGTAACACATGCCCA 2698
Db      9851  CAACCTCTCTCTGATGCTGAGGCTTCTGTAGATATGTTGTAACACATGCCCA 9910
Qy      2699  TGTATATACATACACAGATATGTAACAGATGTACACATGCTCCAGCCAGCTT 2758
Db      9911  TGTATATACATACACAGATATGTAACAGATGTACACATGCTCCAGCCAGCTT 9970
Qy      2759  GCATACCTGACCTGACCCCAAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2818
Db      9971  GCATACCTGACCTGACCCCAAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10030
Qy      2819  CTCACACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2878
Db      10031  CTCACACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10090
Qy      2879  TACCTTGGTGAACCGGTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2938
Db      10091  TACCTTGGTGAACCGGTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10150
Qy      2939  TCTCCGTGACATATCAATTAAGTGGTTGTTACAA 2978

```

```

Db      10151  TCTCCGTGACATATCAATTAAGTGGTTGTTACAA 10190
RESULT 8
AAZ36913
ID      AAZ36913 standard; DNA, 837 BP.
AC      AAZ36913;
XX      13-MAR-2000 (first entry)
DE      DNA encoding a homologue of activator of G protein signalling AGS1.
KW      Activator of G protein signalling; AGS; ras-related G protein;
KW      GTP hydrolysis; G protein activity; phenome response pathway;
KW      G protein-coupled signal transduction; G-gamma selectivity;
KW      cellular signal transduction; AGS1 homologue; ss.
OS      Homo sapiens.
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      1..837
FT      /tag= a
FT      /product= "AGS1 homologue"
MO9958670-A1.
PD      18-NOV-1999.
PF      07-MAY-1999; 99WO-US010151.
PR      08-MAY-1998; 98US-0084842P.
PR      07-OCT-1998; 98US-0103355P.
XX      (CADU-) CADUS PHARM CORP.
XX      Cismowski M, Duzic E;
XX      MPI, 2000-072337/06.
XX      P-98DB; AMY53923.
XX      A new activator of G protein signalling used to treat disorders
XX      characterized by an aberrant AGS protein activity.
XX      Example 15; Page 143-144; 162pp; English.
PS      The present sequence encodes a homologue of a human AGS1 (activator of G
CC      protein signalling (AGS)) protein. The AGS cDNA sequence was isolated
CC      from a human liver cDNA library. The AGS protein exhibits homology to ras
CC      -related G proteins, and contains alterations in conserved amino acids
CC      consistent with a deficiency in GTP hydrolysis activity. AGS stimulates G
CC      protein activity, G protein-coupled signal transduction and the phenome
CC      response pathway in a receptor-independent manner. The AGS protein also
CC      shows G-gamma selectivity, as measured by growth assays in yeast
CC      expressing various mammalian G-gamma constructs, and tissue-specific
CC      expression, as measured by Northern blot analysis. The AGS protein can be
CC      used to screen for compounds that modulate cellular signal transduction.
CC      The protein is used to treat disorders characterized by an aberrant AGS
CC      protein activity or AGS nucleic acid expression
SQ      Sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 U; 0 Other;
Query Match      24.7%; Score 762; DB 3; Length 837;
Best Local Similarity 99.9%; Pred. No. 2e-236;
Matches 812; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      199  CAGCCCGAGCCATGATGAGACTTTGTCAGCGGAACTGACGCTCAGTGTGCCGCC 258
Db      25  CAGCCCGAGCCATGATGAGACTTTGTCAGCGGAACTGACGCTCAGTGTGCCGCC 84
Qy      259  AAAAATCATACCGCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
Db      85  AAAAATCATACCGCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 144

```

```
QY 319 TCTGGCTTCTCAATGCGCGCTTTGAGGACAGATGACACCCACCATCGAGAGCTTCAC 378
DB 145 TCTGCTTCTCAATGCGCGCTTTGAGGACAGATGACACCCACCATCGAGAGCTTCAC 204
QY 379 CGTAAGGATATCAACATCCGCGGCGACATGTACAGAGCTGACATCTTGATACCTTGGC 438
DB 205 CGTAAGGATATCAACATCCGCGGCGACATGTACAGAGCTGACATCTTGATACCTTGGC 264
QY 439 AACCAACCCCTTCCCGCCGATGCGAGGCTGTCATCTCCACAGGGAGATGTTTCACTCG 498
DB 265 AACCAACCCCTTCCCGCCGATGCGAGGCTGTCATCTCCACAGGGAGATGTTTCACTCG 324
QY 499 GTGTTTCAAGCTGATATACCGGAGTCTCTTCATGATGATGATGATGATGATGATGATG 558
DB 325 GTGTTTCAAGCTGATATACCGGAGTCTCTTCATGATGATGATGATGATGATGATGATG 384
QY 559 CTGAGAGGTCAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 618
DB 385 CTGAGAGGTCAGTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 444
QY 619 ATCTGTGACACAGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 678
DB 445 ATCTGTGACACAGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 504
QY 679 GAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 738
DB 505 GAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 564
QY 739 ACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 798
DB 565 ACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 624
QY 799 AGCCCGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
DB 625 AGCCCGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 684
QY 859 TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 918
DB 685 TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 744
QY 919 CGCCCGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 978
DB 745 CGCCCGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 804
QY 979 GCGCCGTGAGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1011
DB 805 GCGCCGTGAGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 837

RESULT 9
AAS90571
ID AAS90571 standard; cDNA, 951 BP.
XX
AC AAS90571;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #26375.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food Supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
```

```
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG26384.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX
XX PT responsible for genetic disorders or other traits and to assess
XX
XX PT biodiversity.
XX
XX
XX Claim 1; SEQ ID NO 26375; 103bp; English.
XX
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX
XX CC sequences. (I) is useful as hybridization probes, polymerase chain
XX
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX
XX CC and in recombinant production of (II). The polynucleotides are also used
XX
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX
XX CC useful for generating antibodies against it, detecting or quantitating a
XX
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX
XX CC involving aberrant protein expression or biological activity. The
XX
XX CC polypeptide and polynucleotide sequences have applications in
XX
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX
XX CC and to produce other types of data and products dependent on DNA and
XX
XX CC amino acid sequences. AAS4197-AAS94564 represent novel human diagnostic
XX
XX CC coding sequences of the invention. Note: The sequence data for this
XX
XX CC patent did not appear in the printed specification, but was obtained in
XX
XX CC electronic format directly from WIPO at
XX
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 951 BP; 202 A; 291 C; 293 G; 165 T; 0 U; 0 Other;
XX
XX
XX Query Match 15.5%; Score 479; DB 5; Length 951;
XX
XX Best Local Similarity 99.8%; Pred. No. 3.3e-145;
XX
XX Matches 529; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX
XX QY 551 AGCAGATCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 610
DB 422 AGCAGATCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481
QY 611 CCATGATCATCTGAGGACAGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 670
DB 482 CCATGATCATCTGAGGACAGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 541
QY 671 CCGAGGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 730
DB 542 CCGAGGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
QY 731 AGAAGAACACCAAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 790
DB 602 AGAAGAACACCAAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
QY 791 ACGAGATGAGCCCGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850
DB 662 ACGAGATGAGCCCGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721
QY 851 CCGAGGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 910
DB 722 CCGAGGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
QY 911 TCGCCCGCCCGCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 970
DB 782 TCGCCCGCCCGCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
QY 971 AAGGCGGAGCCGTGAGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030
```

DB 842 AAGGCCAGGCCCGTGAAGGAGCAATGACCATCCAGTGAAGGAGGATGCTGGGGCGG 901

QY 1031 GGGTTGGCCAGTGCCTTCAAGGAGGTGGCCCCAGATGCCACTGTGCGCA 1080

DB 902 GGGTTGGCCAGTGCCTTCAAGGAGGTGGCCCCAGATGCCACTGTGCGCA 951

RESULT 10

ACH15529

ID ACH15529 standard; cDNA; 469 BP.

XX ACH15529;

AC ACH15529;

XX

DT 13-OCT-2003 (first entry)

XX

DE Human adult brain cDNA #2741.

XX

KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

KW genome mapping; biodiversity; genetic disorder.

XX

OS Homo sapiens.

XX

PN US2003073623-A1.

XX

PD 17-APR-2003.

XX

PF 30-JUL-2001; 2001US-00918995.

XX

PR 30-JUL-2001; 2001US-00918995.

XX

PA (DRMA/) DRMANAC R T.

PA (LABA/) LABAT I.

PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.

PA (JONE/) JONES L W.

XX

PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX

DR WPI; 2003-615964/58.

XX

XX New polynucleotide sequences obtained from various cDNA libraries, useful

PT as hybridization probes, as oligomers for PCR, for chromosome and gene

PT mapping, in the recombinant production of protein, or in generating

PT antisense DNA or RNA.

XX

PS Claim 1; SEQ ID NO 2741; 44pp; English.

XX

CC The invention relates to an isolated polynucleotide comprising any one of

CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was

CC determined by the technique of SBH (sequencing by hybridisation). Also

CC included is a purified polypeptide comprising a sequence corresponding to

CC a reading frame of the novel polynucleotide. The nucleic acid sequences

CC are useful in diagnostics as expressed sequence tags (EST) for

CC identifying expressed genes or for physical mapping of the human genome,

CC in forensics, in assessing biodiversity, or in identifying mutations

CC responsible for genetic disorders and other traits. The nucleotide

CC sequences are also useful as hybridisation probes, as oligomers for PCR,

CC for chromosome and gene mapping, in the recombinant production of

CC protein, or in generating antisense DNA or RNA. The purified polypeptide

CC is useful for generating antibodies specific for it. The present sequence

CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data

CC for this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX

XX Sequence 469 BP; 102 A; 145 C; 128 G; 87 T; 0 U; 7 Other;

XX

QY Query Match 13.8%; Score 424; DB 8; Length 469;

QY Best Local Similarity 100.0%; Pred. No. 2.1e-127;

QY Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1964 GCTCCATTTCACAGAGATACACCCGAGCTCGAGGGGACAGATGACACAGGCCC 2023

XX

DB 45 GCTCCATTTCACAGAGATACACCCGAGCTCGAGGGGACAGATGACACAGGCCC 104

QY 2024 CACCTTTGGTCCCTTAGCAAAATTCAAGGTACAGCTCCACTAGAAACAGAGCTGCCCTTAC 2083

DB 105 CACCTTTGGTCCCTTAGCAAAATTCAAGGTACAGCTCCACTAGAAACAGAGCTGCCCTTAC 164

QY 2084 TGTGCTGTTCTCTCAAGCATTTATTAGACACTTACGTGGTGTGCTGTGCTCTTA 2143

DB 165 TGTGCTGTTCTCTCAAGCATTTATTAGACACTTACGTGGTGTGCTGTGCTCTTA 224

QY 2144 GGAACCAAGAGGGTCCCAAGTCTGAGCCGCGCCGCTGCGCCCAACACTTCT 2203

DB 225 GGAACCAAGAGGGTCCCAAGTCTGAGCCGCGCCGCTGCGCCCAACACTTCT 284

QY 2204 GCACACACAGCGGTGGGAGGCGGGAGAGACAGCTGAGAACCAACTGAGCTGGAG 2263

DB 285 GCACACACAGCGGTGGGAGGCGGGAGAGACAGCTGAGAACCAACTGAGCTGGAG 344

QY 2264 GGATCCGACAGAAAAAGCTCAGGGGCGGTCTTCTCTTGTGCCGGGATTGGGCTATGCTG 2323

DB 345 GGATCCGACAGAAAAAGCTCAGGGGCGGTCTTCTCTTGTGCCGGGATTGGGCTATGCTG 404

QY 2324 GGATCCGACAGAAAAAGCTCAGGGGCGGTCTTCTCTTGTGCCGGGATTGGGCTATGCTG 2383

DB 405 GGATCCGACAGAAAAAGCTCAGGGGCGGTCTTCTCTTGTGCCGGGATTGGGCTATGCTG 464

QY 2384 TCAG 2387

DB 465 TCAG 468

RESULT 11

ACD93776

ID ACD93776 standard; cDNA; 607 BP.

XX

XX ACD93776;

AC ACD93776;

XX

DT 23-SEP-2003 (first entry)

XX

DE Human colon cancer cell expressed cDNA #2188.

XX

KW Open reading frame detection; genome sequencing; colon cancer;

KW breast cancer; population genome analysis; genetic shift; cancer;

KW antibiotic resistance; antibiotic non-tolerance; congenital disease;

KW agriculture; food crop genome; resistance gene; retrovirus;

KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;

KW gene; ss.

XX

OS Homo sapiens.

XX

PN US2002155438-A1.

XX

PD 24-OCT-2002.

XX

PF 27-SEP-1999; 99US-00406117.

XX

PR 20-NOV-1998; 98US-00196716.

XX

PA (SIMP/) SIMPSON A J G.

PA (NETO/) NETO E D.

PA (BREN/) BRENTANI R R.

XX

PI Simpson AJG, Neto ED, Brentani RR;

XX

DR WPI; 2003-182626/18.

XX

PT Determining open reading frames of genome of an organism e.g. a human

PT suffering from cancer involves use of single oligonucleotide primer at

PT low stringency for preparing single-stranded cDNA from mRNA of

PT individual.

XX

PS Example 9; Page 336; 959pp; English.

XX

CC The invention describes a method of determining open reading frames in  
CC the genome of organism, comprising contacting mRNA from cell of organism  
CC with a single oligonucleotide primer (1) at low stringency, preparing  
CC single-stranded cDNA by reverse transcribing mRNA with (1), amplifying  
CC cDNA, sequencing the product, and repeating the contacting, preparing  
CC and amplifying steps with different primers and sequencing resulting  
CC nucleic acids. The method is useful for determining that a known  
CC nucleotide sequence from a genome of an organism corresponds to a  
CC nucleotide sequence from an open reading frame, for preparing a contig,  
CC nucleic acid molecule from a genome of an organism, and for sequencing  
CC all or part of a genome of an organism. mRNA is obtained from mammalian  
CC or human cell which is associated with a pathological condition e.g. a  
CC colon cancer or breast cancer cell. The method is useful for analyses of  
CC populations of subjects and can be used to carry out genetic analyses of  
CC large or small populations. further, it can be used to study living  
CC systems to determine if, e.g. there have been genetic shifts which render  
CC an individual or population more or less likely to be afflicted with  
CC diseases such as cancer, to determine antibiotic resistance or non-  
CC tolerance, and so forth. The method can also be used in the study of  
CC congenital diseases, and the risk of affliction to a foetus, as well as  
CC the study of whether the conditions are likely to be passed to offspring  
CC through ova or sperm. The analyses for pathological conditions can be  
CC carried out in all animals, plants, birds, fish, etc. Using this method,  
CC in the area of agriculture, for example the genomes of food crops can be  
CC studied to determine if resistance genes are present, defects in plant  
CC genomes can also be studied in this way. Similarly, the method permits  
CC determination of the pathogens which integrate into the genome, such as  
CC retroviruses and other integrating viruses such as influenza virus, have  
CC undergone shifts or mutations, which may require different approaches to  
CC therapy. This method is also applied to eukaryotic pathogens, such as  
CC trypanosomes, different types of Plasmodium, etc. The method essentially  
CC eliminates sequencing of non-coding portions. This sequence represents a  
CC polynucleotide isolated from human colon cancer cell cDNA library

XX Sequence 607 BP; 140 A; 149 C; 143 G; 160 T; 0 U; 15 Other;

XX Query Match 12.6%; Score 388; DB 7; Length 607;

Best Local Similarity 100.0%; Pred. No. 8.1e-116; Mismatches 0; Gaps 0;

Matches 388; Conservative 0; Indels 0; Gaps 0;

QY 2229 GAGGAGAGCTGGGAGCCAGAGCTGAGCTGGAGAGGATCCGACAGAAAGCTCAGGCG 2288

DB 12 GAGGAGAGCTGGGAGCCAGAGCTGAGCTGGAGAGGATCCGACAGAAAGCTCAGGCG 71

QY 2289 GGTCTTCTCTTGTGCGCGGATTTGGGCTATGCTGGTACCAACATGTACTCAGGCATGG 2348

DB 72 GGTCTTCTCTTGTGCGCGGATTTGGGCTATGCTGGTACCAACATGTACTCAGGCATGG 131

QY 2349 TGGGTTTGAACCCATAAACCAGGCTTGTATCATGCTTTAAACAAGTATTTTGT 2408

DB 132 TGGGTTTGAACCCATAAACCAGGCTTGTATCATGCTTTAAACAAGTATTTTGT 191

QY 2409 ATTAAATCTCTTAACATATTTGAAGTTTGAAGGCTTAAAGAACTTAGTATCTTCT 2468

DB 192 ATTAAATCTCTTAACATATTTGAAGTTTGAAGGCTTAAAGAACTTAGTATCTTCT 251

QY 2469 ATTGGGCTTCTTGAAGTTTGAAGGCTTAAAGAACTTAGTATCTTCT 2528

DB 252 ATTGGGCTTCTTGAAGTTTGAAGGCTTAAAGAACTTAGTATCTTCT 311

QY 2529 TGTGGGTGGGAGAAAGCTAGAGGCTGGGCAATTCAGTACATCCAGATGTGCTCCTC 2588

DB 312 TGTGGGTGGGAGAAAGCTAGAGGCTGGGCAATTCAGTACATCCAGATGTGCTCCTC 371

QY 2589 TCTTGATGCTTGGCCCTGGGGCCTTCA 2616

DB 372 TCTTGATGCTTGGCCCTGGGGCCTTCA 399

RESULT 12  
ID ACH42415 standard; cDNA, 567 BP.

XX

AC ACH42415;

XX 13-OCT-2003 (first entry)

XX Human foetal brain cDNA #3782.

XX Human; ss; sequencing by hybridisation; SSH; expressed sequence tag; EST;

XX genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUN-2001; 2001US-00918995.

XX 30-JUN-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful

XX as hybridization probes, as oligomers for PCR, for chromosome and gene

XX mapping, in the recombinant production of protein, or in generating

XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 29627; 44p; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
XX 38043 cDNA sequences, appearing as ACH42789-ACH50831, whose sequence was  
XX determined by the technique of SSH (sequencing by hybridisation). Also  
XX included is a purified polypeptide comprising a sequence corresponding to  
XX a reading frame of the novel polynucleotide. The nucleic acid sequences  
XX are useful in diagnostics as expressed sequence tags (EST) for  
XX identifying expressed genes or for physical mapping of the human genome,  
XX in forensics, in assessing biodiversity, or in identifying mutations,  
XX responsible for genetic disorders and other traits. The nucleotide  
XX sequences are also useful as hybridisation probes, as oligomers for PCR,  
XX for chromosome and gene mapping, in the recombinant production of  
XX protein, or in generating antisense DNA or RNA. The purified polypeptide  
XX is useful for generating antibodies specific for it. The present sequence  
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
XX for this patent did not form part of the printed specification, but was  
XX obtained in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html?docid=20030073623

XX Sequence 567 BP; 138 A; 134 C; 147 G; 146 T; 0 U; 2 Other;

XX Query Match 11.6%; Score 358; DB 8; Length 567;

Best Local Similarity 99.8%; Pred. No. 3.8e-106; Mismatches 1; Indels 0; Gaps 0;

Matches 408; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2249 AACTGAGCTGGGAGGATCCGACAGAAAGCTCAGGCGGCTTCTCTTGTGCGCGG 2308

DB 74 AACTGAGCTGGGAGGATCCGACAGAAAGCTCAGGCGGCTTCTCTTGTGCGCGG 133

QY 2309 GATTGGGCTATGCTGGGTACACCATGTACTCAGGAGTGGGTTTGAACCATTAAC 2368

DB 134 GATTGGGCTATGCTGGGTACACCATGTACTCAGGAGTGGGTTTGAACCATTAAC 193

QY 2369 CAAAGGCCCTTGTATCATGCTTTAAACAAGTATTTTGAATCTCTTAACAT 2428

DB 194 CAAAGGCCCTTGTATCATGCTTTAAACAAGTATTTTGAATCTCTTAACAT 253

QY 2429 ATTGAAGTTTGAAGGCTTAAAGAACTTAGTATCTTCTTGTGCGCTTCTGAGTTC 2488

```

Db      ||| 254 ATTTAGTTTATGGCCCTTAAGAACTTAGATCTTCTATTTGGTCTTTCTGAGTTC 313
Qy      ||| 249 AGAGGGGTAACTACTTCTCTCAGGTACACAGCAAGTCTGTGGTGGCAGAAACAGC 2548
Db      ||| 314 AGAGGGGTAACTACTTCTCTCAGGTACACAGCAAGTCTGTGGTGGCAGAAACAGC 373
Qy      ||| 2549 TAGCGCTGGGCACTAGTACATACACAGATGTCCTCTCTTATGATGTCGCCCTGG 2608
Db      ||| 374 TAGCGCTGGGCACTAGTACATACACAGATGTCCTCTCTTATGATGTCGCCCTGG 433
Qy      ||| 2609 GGCCTTCAGGGCTTTGGGACATCTTGTCTCAACCCCTCTCCCTAGATCA 2657
Db      ||| 434 GGCCTTCAGGGCTTTGGGACATCTTGTCTCAACCCCTCTCCCTAGATCA 482

```

## RESULT 13

AA124259 ID AA124259 standard; DNA; 368 BP.

XX AC AA124259;

XX DT 12-OCT-2001 (first entry)

XX DE Probe #14192 for gene expression analysis in human cervical cell sample.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX KM cervical cancer; ss.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000670.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PS Claim 25; SEQ ID NO 14192; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes (SENPs). The present sequence is one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably cervical cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 368 BP; 73 A; 119 C; 103 G; 73 T; 0 U; 0 Other;

Query Match 9.2%; Score 283; DB 4; Length 368;  
Best Local Similarity 100.0%; Pred. No. 6.4e-82;  
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      ||| 199 CAGCCCGAGCCATGATGAACCTTTGTCAGCGGGAACCTGACAGCTGATGCCGCC 258
Db      ||| 86 CAGCCCGAGCCATGATGAACCTTTGTCAGCGGGAACCTGACAGCTGATGCCGCC 145
Qy      ||| 259 AAAAATCATACCCGATGCTGTGTGCTCTCTCTGCGGTGGCAAGAGCTTCATCGTG 318
Db      ||| 146 AAAAATCATACCCGATGCTGTGTGCTCTCTCTGCGGTGGCAAGAGCTTCATCGTG 205
Qy      ||| 319 TCTGCTTCTCTCATGAGCGCGCTTTGAGGACAGATACACCCACATCGAGAGACTTCAC 378
Db      ||| 206 TCTGCTTCTCTCATGAGCGCGCTTTGAGGACAGATACACCCACATCGAGAGACTTCAC 265
Qy      ||| 379 CGTAAGGTATACAACTATCCGCGGACATGTATACAGCTCGACATCTGATATCTTGGC 438
Db      ||| 266 CGTAAGGTATACAACTATCCGCGGACATGTATACAGCTCGACATCTTGGATACCTCTGGC 325
Qy      ||| 439 AACCACTCTTCCCGCATGCGAGGCTGTTCATCTTCACAG 481
Db      ||| 326 AACCACTCTTCCCGCATGCGAGGCTGTTCATCTTCACAG 368

```

## RESULT 14

ABA69388 ID ABA69388 standard; DNA; 368 BP.

XX AC ABA69388;

XX DT 01-FEB-2002 (first entry)

XX DE Human foetal liver single exon nucleic acid probe #17693.

XX KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PS Claim 4; SEQ ID NO 17693; 639bp + Sequence Listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 368 BP; 73 A; 119 C; 103 G; 73 T; 0 U; 0 Other;

Query Match 9.2%; Score 283; DB 4; Length 368;  
Best Local Similarity 100.0%; Pred. No. 6.4e-82;

Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CAGCCCCGAGCATGATGAAGACTTTGTCCAGCGGGAATGCAAGCTCAAGTGTCCGCC 258  
DB 86 CAGCCCCGAGCATGATGAAGACTTTGTCCAGCGGGAATGCAAGCTCAAGTGTCCGCC 145

QY 259 AAAAATCATACCGCATGATGATGCTGTGCTCTCTCGGGTGGGCAAGAGCTCCATCGTG 318  
DB 146 AAAAATCATACCGCATGATGATGCTGTGCTCTCTCGGGTGGGCAAGAGCTCCATCGTG 205

QY 319 TCTGCTTCTCAATGAGCGCGCTTTGAGGACCAATGACACCCACCATCGAGACTTCAC 378  
DB 206 TCTGCTTCTCAATGAGCGCGCTTTGAGGACCAATGACACCCACCATCGAGACTTCAC 265

QY 379 CGTAAGGTATACCAATCCGCGGACATGATACCAAGCTCGACATCTCGATCTTGGC 438  
DB 266 CGTAAGGTATACCAATCCGCGGACATGATACCAAGCTCGACATCTCGATCTTGGC 325

QY 439 AACCACTTCCCGCATGCGGAGGCTGTCCATCTTCACAG 481  
DB 326 AACCACTTCCCGCATGCGGAGGCTGTCCATCTTCACAG 368

RESULT 15  
AA149541  
ID AA149541 standard; DNA; 368 BP.

AC AA149541;

DT 17-OCT-2001 (first entry)

DE Probe #18227 used to measure gene expression in human placenta sample.

KM Probe; microarray; human; placenta; antenatal diagnosis;  
genetic disorder; ss.

OS Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-0063236P.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024283.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing

PS gene expression in human placenta.

PS Claim 25; SEQ ID NO 18227; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENPs).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders

SQ Sequence 368 BP; 73 A; 119 C; 103 G; 73 T; 0 U; 0 Other;

Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CAGCCCCGAGCATGATGAAGACTTTGTCCAGCGGGAATGCAAGCTCAAGTGTCCGCC 258  
DB 86 CAGCCCCGAGCATGATGAAGACTTTGTCCAGCGGGAATGCAAGCTCAAGTGTCCGCC 145

QY 259 AAAAATCATACCGCATGATGATGCTGTGCTCTCTCGGGTGGGCAAGAGCTCCATCGTG 318  
DB 146 AAAAATCATACCGCATGATGATGCTGTGCTCTCTCGGGTGGGCAAGAGCTCCATCGTG 205

QY 319 TCTGCTTCTCAATGAGCGCGCTTTGAGGACCAATGACACCCACCATCGAGACTTCAC 378  
DB 206 TCTGCTTCTCAATGAGCGCGCTTTGAGGACCAATGACACCCACCATCGAGACTTCAC 265

QY 379 CGTAAGGTATACCAATCCGCGGACATGATACCAAGCTCGACATCTCGATCTTGGC 438  
DB 266 CGTAAGGTATACCAATCCGCGGACATGATACCAAGCTCGACATCTCGATCTTGGC 325

QY 439 AACCACTTCCCGCATGCGGAGGCTGTCCATCTTCACAG 481  
DB 326 AACCACTTCCCGCATGCGGAGGCTGTCCATCTTCACAG 368

Search completed: June 20, 2004, 12:14:22  
Job time : 1147 secs

THIS PAGE BLANK (USPTO)





JOURNAL Patent: WO 02062849-A 1 15-AUG-2002;  
PE Corporation (NY) (US)  
FEATURES Location/Qualifiers  
SOURCE 1. .3082  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 3082; DB 6; Length 3082;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCTGCGCGCGGAGAAAGACCGGCGCCAGCGCGCTCCGAGGAGCGGAGG 60  
DB 1 GGGCTGCGCGCGGAGAAAGACCGGCGCCAGCGCGCTCCGAGGAGCGGAGG 60  
QY 61 GAGGATCCCGCGAGTAGACCCGAGGAGCACACAGACTTGGAGAGGCTCGCGCTGGA 120  
DB 61 GAGGATCCCGCGAGTAGACCCGAGGAGCACACAGACTTGGAGAGGCTCGCGCTGGA 120  
QY 121 GCAGCAGGAGACTCCCGGAGCTCCGCGCTTCAGGAGAGCTCTGAGCCGTGCGA 180  
DB 121 GCAGCAGGAGACTCCCGGAGCTCCGCGCTTCAGGAGAGCTCTGAGCCGTGCGA 180  
QY 181 GGGCGGCGCGCAATTCAGACCCCGAGCAATGATGAAGCTTTCAGGCGGAGACTGC 240  
DB 181 GGGCGGCGCGCAATTCAGACCCCGAGCAATGATGAAGCTTTCAGGCGGAGACTGC 240  
QY 241 ACGCTCAGTGTGCCCGCAAAACTCATACCGCATGTGTGTGTGTGTGTGTGTGTGT 300  
DB 241 ACGCTCAGTGTGCCCGCAAAACTCATACCGCATGTGTGTGTGTGTGTGTGTGTGT 300  
QY 301 GGGAGAGCTCAGT 360  
DB 301 GGGAGAGCTCAGT 360  
QY 361 ACCATCGAGACTTCACCGTAGATATCAACATCCGCGGAGCATATGACAGCTGAC 420  
DB 361 ACCATCGAGACTTCACCGTAGATATCAACATCCGCGGAGCATATGACAGCTGAC 420  
QY 421 ATCTGTGATACCTCTGGCAACACCCCTTCCCGCATGCGAGGCTGTCCATCTCA 480  
DB 421 ATCTGTGATACCTCTGGCAACACCCCTTCCCGCATGCGAGGCTGTCCATCTCA 480  
QY 481 GGGGATGTCTTCACTGT 540  
DB 481 GGGGATGTCTTCACTGT 540  
QY 541 CGCCTTGAAGAGAGATCTGTGAGGTCAAGTCTGCTGAAGAACAAAGAGAGGCG 600  
DB 541 CGCCTTGAAGAGAGATCTGTGAGGTCAAGTCTGCTGAAGAACAAAGAGAGGCG 600  
QY 601 GCGGAGCTGCCATGT 660  
DB 601 GCGGAGCTGCCATGT 660  
QY 661 GTGCGCCACACCGAGGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720  
DB 661 GTGCGCCACACCGAGGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720  
QY 721 GTGTGCGCCAGAGAAACACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
DB 721 GTGTGCGCCAGAGAAACACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
QY 781 AAGTGTGCAACAGATGAGCCCGCTGTGATGCAAGATCTCCGTGACAGTGTGAC 840  
DB 781 AAGTGTGCAACAGATGAGCCCGCTGTGATGCAAGATCTCCGTGACAGTGTGAC 840  
QY 841 GCGCTTCAACCGAGGCGCTTGTGATGCGCGGTCAAGAGAGATGAGCGCTATGAGATG 900  
DB 841 GCGCTTCAACCGAGGCGCTTGTGATGCGCGGTCAAGAGAGATGAGCGCTATGAGATG 900

QY 901 GTTCGCGCTTCCG 960  
DB 901 GTTCGCGCTTCCG 960  
QY 961 GTTCCTCGGAGAGCG 1020  
DB 961 GTTCCTCGGAGAGCG 1020  
QY 1021 GCTGGGCGCGGCGCTTGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080  
DB 1021 GCTGGGCGCGGCGCTTGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080  
QY 1081 TCTCCCGACCGAGGCG 1140  
DB 1081 TCTCCCGACCGAGGCG 1140  
QY 1141 CCGCGGCGCGTGGCGCTCCGCAATTCGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1200  
DB 1141 CCGCGGCGCGTGGCGCTCCGCAATTCGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1200  
QY 1201 GTCCAGACTCTCTGT 1260  
DB 1201 GTCCAGACTCTCTGT 1260  
QY 1261 GAGTTAGCAGAGACTCAAGTTACACCTTCCTCTGCGGAGTGAAGAAATGTGATGC 1320  
DB 1261 GAGTTAGCAGAGACTCAAGTTACACCTTCCTCTGCGGAGTGAAGAAATGTGATGC 1320  
QY 1321 CAGAGGAGTGAAGATTCTGT 1380  
DB 1321 CAGAGGAGTGAAGATTCTGT 1380  
QY 1381 GGAACAGAGAGCGCAAGTAGAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440  
DB 1381 GGAACAGAGAGCGCAAGTAGAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440  
QY 1441 TGGGTGCGAGCTGGAGAACTTCTCTCCAGCCCTGTCAATCTTACGCTCTGTGTACGT 1500  
DB 1441 TGGGTGCGAGCTGGAGAACTTCTCTCTCCAGCCCTGTCAATCTTACGCTCTGTGTACGT 1500  
QY 1501 GCTCTGCAACCCCTTCCACCCCGAGCAACACAAAGTTGGCCCGAGCTGCGCTGAC 1560  
DB 1501 GCTCTGCAACCCCTTCCACCCCGAGCAACACAAAGTTGGCCCGAGCTGCGCTGAC 1560  
QY 1561 ATTGAGCAGTGAATCTGT 1620  
DB 1561 ATTGAGCAGTGAATCTGT 1620  
QY 1621 CCACTTGAACGAGCCCACTCTGTGACCGGTTCTGTAGGCTCTCTCTAGGTCCCTCC 1680  
DB 1621 CCACTTGAACGAGCCCACTCTGTGACCGGTTCTGTAGGCTCTCTCTAGGTCCCTCC 1680  
QY 1681 GCGCGAGAGTTGT 1740  
DB 1681 GCGCGAGAGTTGT 1740  
QY 1741 ATGGAATCAATTTGATCTGTAAAGCTGTGACTCTCTCTGTGCGAGGCGCTCACCCAG 1800  
DB 1741 ATGGAATCAATTTGATCTGTAAAGCTGTGACTCTCTCTGTGCGAGGCGCTCACCCAG 1800  
QY 1801 TTCAAGTCAACGAGCTTCAACCGGAGAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
DB 1801 TTCAAGTCAACGAGCTTCAACCGGAGAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
QY 1861 TGGCTGT 1920  
DB 1861 TGGCTGT 1920  
QY 1921 GCAAGAGCGGAGTGGGAGGCGAGGAGCTTTTGTAAAGTCTCTCTCTCTCTCTCTCTCTCT 1980  
DB 1921 GCAAGAGCGGAGTGGGAGGCGAGGAGCTTTTGTAAAGTCTCTCTCTCTCTCTCTCTCTCT 1980  
QY 1981 GATACACCGAGACTGGAGGAGCAAGTAGACACAGGCGCCACCTTGTCTCCCTAGC 2040

[illegible]

Db 3061 AAAAAAAAAAAAAAAAAAAAAA 3082

RESULT 2	
LOCUS	BC013419
DEFINITION	Homo sapiens RASD family, member 2, mRNA (cdna clone MGC:4834
ACCESSION	BC013419
VERSION	BC013419.2
KEYWORDS	IMAGE:3604196), complete cds.
SOURCE	MGc.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	1 (bases 1 to 3049)
	1 (bases 1 to 3049)
REFERENCE	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
AUTHORS	Klausner,R.D., Collins,F.S., Wagner,K.H., Shennan,C.M., Schuler,G.D.,
	Altshuler,S.P., Zeeberg,B., Bietow,K.H., Schaefer,C.F., Bhat,N.K.,
	Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heien,F.,
	Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
	Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
	Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
	Canciuci,P., Prange,C., Raha,S.S., Locuelli,N.A., Peters,G.J.,
	Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
	McEranan,K.U., Malek,U.A., Gunaratne,P.H., Richards,S.,
	Wolter,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
	Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
	Fahney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
	Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
	Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
	Dickson,S.C., Rodriguez,A.C., Grimmwood,J., Schmutz,U., Myers,R.M.,
	Buchterfeld,Y.S., Krzywnski,M.I., Skalska,U., Smalish,D.E.,
	Schmerch,A., Schein,J.E., Jones,S.J., and Marx,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length
	human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)
MEDLINE	22388257
PUBMED	12477932
REFERENCE	2 (bases 1 to 3049)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (31-AUG-2001) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	On Aug 19, 2003 this sequence version replaced gi:15426590.
	Contact: MGC help desk
	Email: <a href="mailto:cgaps-rc@mail.nih.gov">cgaps-rc@mail.nih.gov</a>
	Tissue Procurement: ATCC
	cDNA Library Preparation: Rubin Laboratory
	DNA Sequencing by: Institute for Systems Biology
	<a href="http://www.systemsbio.org">http://www.systemsbio.org</a>
	contact: <a href="mailto:amdan@systemsbiology.org">amdan@systemsbiology.org</a>
	Anup Madan, Jessica Fahney, Erin Helton, Mark Kettman, Anuradha
	Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LINL at: <http://image.lim.gov>  
Series: IRAL Plate: 11 Row: B Column: 18  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 22027485.

FEATURES

Source

1..3049

location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:4834 IMAGE:3604196"

/tissue\_type="uterus, endometrium adenocarcinoma"

/clone\_lib="NIH\_MGC\_44"

/lab\_host="DH10B-R"

/note="Vector: POTB7"  
1. 3049  
/gene="RASD2"  
/note="Synonyms: TEN2, MGC:4834, Rhes"  
/db\_xref="LocusID:23551"  
208.1008  
/product="RASD family, member 2"  
/codon\_start=1  
/protein\_id="AAH13419.1"  
/db\_xref="GI:15426591"  
/db\_xref="LocusID:23551"  
/translation="MKKTLSSGNCITLSPANKSYRMVVLGASRVKKSIVSRFLNGR  
EDQYTPLEDPRKRVNIRGDMYOLDIDTSNHPFAMRRLSLTGVFLVSLN  
RESFDEKRLQKOLLEVKSCIKNKKRAALPMVTCGNKHGSLQVPTAEALIV  
SGDENCAFYFSAKNTNVDPMFYVLSMALKPHMBPALHKKISVQYGAFHRPRPC  
MRVKEMDAYGMVSPFARFSPVNSDLKIKAKVIREQARERDCTIQ"  
misc\_feature  
259..780  
/note="RAS; Region: Ras subfamily of RAS small GTPases"  
/db\_xref="CDD:smart00173"

## ORIGIN

Query Match 94.9%; Score 2925; DB 9; Length 3049;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3025; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 26 GCCGCGCCAGCCCGCGCTCCGAGCAGCGAGGAGAGATCCCGCGCAGTGAACCCGG 85  
Db 23 GCCGCGCCAGCCCGCGCTCCGAGCAGCGAGGAGAGATCCCGCGCAGTGAACCCGG 82  
QY 86 AGCCACCAAGACTCTGGAGAGCTCGCGCGCTGAGCAGCAGAGCTCCCGCAGCTCC 145  
Db 83 AGCCACCAAGACTCTGGAGAGCTCGCGCGCTGAGCAGCAGAGCTCCCGCAGCTCC 142  
QY 146 CGGCGCTTCCAGGAGCTCTCTGAGCCGTGCGAAGGCCCGCGCCGATTCGACGCC 205  
Db 143 CGGCGCTTCCAGGAGCTCTCTGAGCCGTGCGAAGGCCCGCGCCGATTCGACGCC 202  
QY 206 GAGCCATGATGAAGCTTTGTCAGCGGGGAACTGAGCGCTAGTGTGCCGCAAAAAC 265  
Db 203 GAGCCATGATGAAGCTTTGTCAGCGGGGAACTGAGCGCTAGTGTGCCGCAAAAAC 262  
QY 266 CATACCGCATGAGTGTGCTGGGTGCTCTCGGGTGGGCAAGAGCTCACTGTGTCTG 325  
Db 263 CATACCGCATGAGTGTGCTGGGTGCTCTCGGGTGGGCAAGAGCTCACTGTGTCTG 322  
QY 326 TCTCTCATATGCGCTTTGAGGACAGATACACACCCATGAGAGCTTCCACCGTAAG 385  
Db 323 TCTCTCATATGCGCTTTGAGGACAGATACACACCCATGAGAGCTTCCACCGTAAG 382  
QY 386 TATACCAATCCGGGGGAGATGTAACAGCTCGAATCTGAGATCTTGGCAACAC 445  
Db 383 TATACCAATCCGGGGGAGATGTAACAGCTCGAATCTGAGATCTTGGCAACAC 442  
QY 446 CTTTCCCGCCATCGCAGAGCTGTCTCATCTCAAGGGAGTGTCTTCACTCTGTGTCA 505  
Db 443 CTTTCCCGCCATCGCAGAGCTGTCTCATCTCAAGGGAGTGTCTTCACTCTGTGTCA 502  
QY 506 GCTCTGATTAACCGGAGTCTTTGATGAGTCAAGCGCTTCAAGAGCATCTCTGAGG 565  
Db 503 GCTCTGATTAACCGGAGTCTTTGATGAGTCAAGCGCTTCAAGAGCATCTCTGAGG 562  
QY 566 TCAAGTCTCTGCTGAAGAACAAACAAAGAGGGGGGAGAGCTGCCCATGTGATCTGTG 625  
Db 563 TCAAGTCTCTGCTGAAGAACAAACAAAGAGGGGGGAGAGCTGCCCATGTGATCTGTG 622  
QY 626 GCAACAAGAACGACACGCGAGCTGTGCGCAGAGTGCACACACCGAGGCGCAGCTGC 685  
Db 623 GCAACAAGAACGACACGCGAGCTGTGCGCAGAGTGCACACACCGAGGCGCAGCTGC 682  
QY 686 TGGTGTGGGGGAGAGAACTCCGCTTACTTGAGAGTGTGGCCAAAGAAACCAACG 745  
Db 683 TGGTGTGGGGGAGAGAACTCCGCTTACTTGAGAGTGTGGCCAAAGAAACCAACG 742

QY 746 TGAAGAGATGTTCTAGTGTCTTTCAGATAGGCCAAGCTGCCACAGATGAGCCCG 805  
Db 743 TGAAGAGATGTTCTAGTGTCTTTCAGATAGGCCAAGCTGCCACAGATGAGCCCG 802  
QY 806 CCTGTGATCGCAAGATCTCCGTGAGTACGAGTGAACGCTTTCACACCAAGGCTTCTGCA 865  
Db 803 CCTGTGATCGCAAGATCTCCGTGAGTACGAGTGAACGCTTTCACACCAAGGCTTCTGCA 862  
QY 866 TGGCGCCGCTCAAGAAATGAGACGCTTATGAGATGATGTCTGCTTGGCCCGCCGCCA 925  
Db 863 TGGCGCCGCTCAAGAAATGAGACGCTTATGAGATGATGTCTGCTTGGCCCGCCGCCA 922  
QY 926 GCGTCAACAGTGAACCTCAAGATCATCAAGGCAAGTCTTCCGGAAGGCCAGGCGCTG 985  
Db 923 GCGTCAACAGTGAACCTCAAGATCATCAAGGCAAGTCTTCCGGAAGGCCAGGCGCTG 982  
QY 986 AGAGGACAAAGTGAACCATTCAGTGAACGAGGATGCTGGGGCGGGCTTGGCCAGTGC 1045  
Db 983 AGAGGACAAAGTGAACCATTCAGTGAACGAGGATGCTGGGGCGGGCTTGGCCAGTGC 1042  
QY 1046 TTCAAGGAGTGGGCCCAAGATGCCCCACTGTGCGATCTTCCCAACCGAGGCCCGGACGA 1105  
Db 1043 TTCAAGGAGTGGGCCCAAGATGCCCCACTGTGCGATCTTCCCAACCGAGGCCCGGACGA 1102  
QY 1106 GTCTTGTTCACAGACTTTAGGACCAAGCTGAGAGCCCGCGGCGCTGGCCTCCGCAAT 1165  
Db 1103 GTCTTGTTCACAGACTTTAGGACCAAGCTGAGAGCCCGCGGCGCTGGCCTCCGCAAT 1162  
QY 1166 TCGTCTCCCTTCTCAGACGCTTTCCTGAGTTCGCTTGTTCACAGCTCTTGTGTTCAT 1225  
Db 1163 TCGTCTCCCTTCTCAGACGCTTTCCTGAGTTCGCTTGTTCACAGCTCTTGTGTTCAT 1222  
QY 1226 CTCTCTGTGGAGAGACATCTCTGAGCTTCAAGATTAGAGAGACTCAAGTTACA 1285  
Db 1223 CTCTCTGTGGAGAGACATCTCTGAGCTTCAAGATTAGAGAGACTCAAGTTACA 1282  
QY 1286 CCTCTCTCTCCTGGGTTGGAAGATGTTGATATCCAGAGGGGTGAAGATTGCTGCTCA 1345  
Db 1283 CCTCTCTCTCCTGGGTTGGAAGATGTTGATATCCAGAGGGGTGAAGATTGCTGCTCA 1342  
QY 1346 TATGAGACCTTCCTGGGACAAAGCTCAGAGATGAAGAAACACAGAGGCCAGATGAAG 1405  
Db 1343 TATGAGACCTTCCTGGGACAAAGCTCAGAGATGAAGAAACACAGAGGCCAGATGAAG 1402  
QY 1406 GTCTCTCTCTCTCGGATTAACACCCAGCTTGTGTTGGGTGGAGCTGGAGAACTTCTC 1465  
Db 1403 GTCTCTCTCTCTCGGATTAACACCCAGCTTGTGTTGGGTGGAGCTGGAGAACTTCTC 1462  
QY 1466 TCCAGGCTCTGCAACTCTTAAGCTCTGATTCAGCTGCTCTGACACCCCTCCACCCCA 1525  
Db 1463 TCCAGGCTCTGCAACTCTTAAGCTCTGATTCAGCTGCTCTGACACCCCTCCACCCCA 1522  
QY 1526 GCACACACAAAGTTGGCCCCCAGCTGCGCTGACATTTAGACAGTGAATCTGTGTCTG 1585  
Db 1523 GCACACACAAAGTTGGCCCCCAGCTGCGCTGACATTTAGACAGTGAATCTGTGTCTG 1582  
QY 1586 AAGGGGGGCTGGCCACACCTCTTGAACCAAGCCCAACCACTTAAGACACAGCCCACTCTG 1645  
Db 1583 AAGGGGGGCTGGCCACACCTCTTGAACCAAGCCCAACCACTTAAGACACAGCCCACTCTG 1642  
QY 1646 ACCGCTTCTCAGCTCTCTCTCTGAGTCCCTCCGCGCCGACAGTGTGCTTGTGTG 1705  
Db 1643 ACCGCTTCTCAGCTCTCTCTCTGAGTCCCTCCGCGCCGACAGTGTGCTTGTGTG 1702  
QY 1706 TTGAGCTGTTTTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765  
Db 1703 TTGAGCTGTTTTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1762  
QY 1766 CTATGATCTCCCTCTGTTGGCAGGCGCTTCAACCCAGTTCATATCAAGGCGCTCCACCCGG 1825  
Db 1763 CTATGATCTCCCTCTGTTGGCAGGCGCTTCAACCCAGTTCATATCAAGGCGCTCCACCCGG 1822  
QY 1826 ACGCTTCTCTCTGCTGCCAAACAGGGTTCGCTGGCTGTGTTGAGCTAGACATTTGA 1885

```

Db      1823  ACGCTTCCTCTGCTGCCAAACAGGGTTTCGGTGGCCCTGTGTGACGTAACATTTAA 1882
Qy      1886  CCTCCGCAATTGAGCTCCACGGTTTACAGACATTTGCAACAAGCTGGGGTGGGCAAGGCA 1945
Db      1883  CCTCCGCAATTGAGCTCCACGGTTTACAGACATTTGCAACAAGCTGGGGTGGGCAAGGCA 1942
Qy      1946  GGAAGCTGTTTTTTTAAATGCTCCCATTTCCACAGAGATCCACGAGCTCGGAGGGGAC 2005
Db      1943  GGAAGCTGTTTTTTTAAATGCTCCCATTTCCACAGAGATCCACGAGCTCGGAGGGGAC 2002
Qy      2006  ACGATGAGACACAGAGCCCACTTTGTCCCTAGCAAAATTCAGGGTACAGCTCCACCTAG 2065
Db      2003  ACGATGAGACACAGAGCCCACTTTGTCCCTAGCAAAATTCAGGGTACAGCTCCACCTAG 2062
Qy      2066  AACCAAGGCTGCTCTACTGTGCTGCTCTCAAGCATTTATTAAGCACTTACGAGTGC 2125
Db      2063  AACCAAGGCTGCTCTACTGTGCTGCTCTCAAGCATTTATTAAGCACTTACGAGTGC 2122
Qy      2126  TGGGTTCACTGTGCTCTAGGAAACCAAGAGGTTCCCAAGTCCCTGCTGCGCCCT 2185
Db      2123  TGGGTTCACTGTGCTCTAGGAAACCAAGAGGTTCCCAAGTCCCTGCTGCGCCCT 2182
Qy      2186  GCTGCCCAACCACTTCTGACACACAGCGGTGGGGAAGAGAGAGAGAGAGAGAGAGAG 2245
Db      2183  GCTGCCCAACCACTTCTGACACACAGCGGTGGGGAAGAGAGAGAGAGAGAGAGAGAG 2242
Qy      2246  CAGAACTGAGCTTGGGAGGAGATCCGACAGAAAGCTCAGGGGAGGTCTTCTCTTGTGCC 2305
Db      2243  CAGAACTGAGCTTGGGAGGAGATCCGACAGAAAGCTCAGGGGAGGTCTTCTCTTGTGCC 2302
Qy      2306  CGGATTTGGGCTAATGCTGGGTACCAACATGTACTCAGGCAATGATGGTTTGAACCCATA 2365
Db      2303  CGGATTTGGGCTAATGCTGGGTACCAACATGTACTCAGGCAATGATGGTTTGAACCCATA 2362
Qy      2366  AACCAAGGCTGCTGATCAGCTCTTAAACAAGATATTTTGTATTTTAACTCTCTAAA 2425
Db      2363  AACCAAGGCTGCTGATCAGCTCTTAAACAAGATATTTTGTATTTTAACTCTCTAAA 2422
Qy      2426  CATATTAAGTTTAAAGGAGCCCTTAAGAACTTATGATCTTCTATTTGGGCTTTTCTAGG 2485
Db      2423  CATATTAAGTTTAAAGGAGCCCTTAAGAACTTATGATCTTCTATTTGGGCTTTTCTAGG 2482
Qy      2486  TTGAGAGAGGTAATGTAATCTTCTCCAGGTACACAGAGATCTGTGGTGGGGAAGCA 2545
Db      2483  TTGAGAGAGGTAATGTAATCTTCTCCAGGTACACAGAGATCTGTGGTGGGGAAGCA 2542
Qy      2546  AGCTAGAGGCTGGGCAATTCAGTAACATCAAGATGAGTCCCTCTTGAATGCTTGGGCC 2605
Db      2543  AGCTAGAGGCTGGGCAATTCAGTAACATCAAGATGAGTCCCTCTTGAATGCTTGGGCC 2602
Qy      2606  TGGGAGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTCTCTCTAGATCAGTCTGTA 2665
Db      2603  TGGGAGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTCTCTCTAGATCAGTCTGTA 2662
Qy      2666  GGGTCCCTGATGATTTGTGTAACACATGCTCATGATATATCAAGTACACACATGTAAC 2725
Db      2663  GGGTCCCTGATGATTTGTGTAACACATGCTCATGATATATCAAGTACACACATGTAAC 2722
Qy      2726  ACAAGATGTAACATGCTCCAGGCCCAAGCTGATACCTGCAACCTGCAACCCAGGCTT 2785
Db      2723  ACAAGATGTAACATGCTCCAGGCCCAAGCTGATACCTGCAACCTGCAACCCAGGCTT 2782
Qy      2786  GGGCCCTGCTGCTGTGTGCTCAAGAGCAGAGCTCAACCTGCTGCTGCTCCCTTCC 2845
Db      2783  GGGCCCTGCTGCTGTGTGCTCAAGAGCAGAGCTCAACCTGCTGCTGCTCCCTTCC 2842
Qy      2846  CACCACTGCTGAGCTTCTGAGAGACACAGATCTTGGCTGACCGGTGTGTGGCC 2905
Db      2843  CACCACTGCTGAGCTTCTGAGAGACACAGATCTTGGCTGACCGGTGTGTGGCC 2902
Qy      2906  GCTCTACCCAGGACAGCCCGCAACATGATCTCCGTGTACATCACTAAAGTG 2965

```

```

Db      2903  GCTCTACCCAGGACAGCCCGCACCATGATCTCCGTACACTATCAATCAATAAGTG 2962
Qy      2966  GGTTCCTTACAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 3025
Db      2963  GGTTCCTTACAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 3022
Qy      3026  AAAAAAAAAAAAAAAAAAAAAAA 3052
Db      3023  AAAAAAAAAAAAAAAAAAAAAAA 3049

RESULT 3
AX393267      2973 bp      DNA      linear      PAT 23-MAR-2002
LOCUS
DEFINITION
Sequence 197 from Patent WO0210217.
ACCESSION
AX393267
VERSION
AX393267.1 GI:19701305
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
St Croix, B., Kinzler, K.W. and Vogelstein, B.
Endothelial cell expression patterns
Patent: WO 0210217-A 197 07-FEB-2002;
The Johns Hopkins University (US)
Location/Qualifiers
1..2973
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source

ORIGIN

Query Match      94.1%; Score 2900; DB 6; Length 2973;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      26  GCGGCGCCAGCCGCGGCTCCGAGCAGCGGAGAGATCCCGCGAGTGAACCGGG 85
Db      23  GCGGCGCCAGCCGCGGCTCCGAGCAGCGGAGAGATCCCGCGAGTGAACCGGG 82
Qy      86  AGCCACACAGACTCTGGAGAGGCTCGGCGGCTGGAGAGAGAGAGAGAGAGAGAGAG 145
Db      83  AGCCACACAGACTCTGGAGAGGCTCGGCGGCTGGAGAGAGAGAGAGAGAGAGAGAG 142
Qy      146  CGGCGCTTCCAGGAGCTCTCTGAGCGGTGCGCAGAGGCGCGCCGCGCATTTCCAGCCC 205
Db      143  CGGCGCTTCCAGGAGCTCTCTGAGCGGTGCGCAGAGGCGCGCCGCGCATTTCCAGCCC 202
Qy      206  GAGCGATGATGAGACTTTTGTCCAGCGGGAATGCAAGCTCAAGTGTCCCGCAAAACT 265
Db      203  GAGCGATGATGAGACTTTTGTCCAGCGGGAATGCAAGCTCAAGTGTCCCGCAAAACT 262
Qy      266  CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
Db      263  CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 322
Qy      326  TCTCATGAGCGCGCTTGAAGAGACAGTACACACCACTGAGAGACTTCCACGTAAG 385
Db      323  TCTCATGAGCGCGCTTGAAGAGACAGTACACACCACTGAGAGACTTCCACGTAAG 382
Qy      386  TATACAAATCCGCGGAGACATGTAACAGTGAATCTGATATCTTGGCAACCAAC 445
Db      383  TATACAAATCCGCGGAGACATGTAACAGTGAATCTGATATCTTGGCAACCAAC 442
Qy      446  CTTTCCCGGAGATCCGAGGCTGCTCACTCTCAAGGGAGATGTTATCTGATCTGATCTCA 505
Db      443  CTTTCCCGGAGATCCGAGGCTGCTCACTCTCAAGGGAGATGTTATCTGATCTGATCTCA 502
Qy      506  GCTGATTAACCGGAGATCTCTGATGAGTCAAGCCCTTTCAGAAACAGATCTGTGAGG 565
Db      503  GCTGATTAACCGGAGATCTCTGATGAGTCAAGCCCTTTCAGAAACAGATCTGTGAGG 562

```

QY 566 TCAAGTCTGCTGAAGAACAAGACCAAGAGGCGCGAGCTGCCCCATGCTCATCTGTG 625  
Db 563 TCAAGTCTGCTGAAGAACAAGACCAAGAGGCGCGAGCTGCCCCATGCTCATCTGTG 622  
QY 626 GCAACAAGAAAGACACCGGAGCTGTGCTCCGACAGGTGCCACCAACGAGGCGGAGCTGC 685  
Db 623 GCAACAAGAAAGACACCGGAGCTGTGCTCCGACAGGTGCCACCAACGAGGCGGAGCTGC 682  
QY 686 TGTGTGTGGGCGACGAGAACTCCGCTTACTTGAGGTGTGCGCCAAAGAAACACCAACG 745  
Db 683 TGTGTGTGGGCGACGAGAACTCCGCTTACTTGAGGTGTGCGCCAAAGAAACACCAACG 742  
QY 746 TGGACGAGATGTTCTAGGTGCTTTACAGATGGCCAACTGCAACAGAGATGAGCCCG 805  
Db 743 TGGACGAGATGTTCTAGGTGCTTTACAGATGGCCAACTGCAACAGAGATGAGCCCG 802  
QY 806 CCTGTGATCGCAAGATCTCCGTGACGTACGCGCTTCCACCCCGAGGCGCTTCTGCA 865  
Db 803 CCTGTGATCGCAAGATCTCCGTGACGTACGCGCTTCCACCCCGAGGCGCTTCTGCA 862  
QY 866 TGGCGCGGCTCAAGAGATGGAACGCTATGCGATGCTTCCGCTTCCGCGCGCGCCCA 925  
Db 863 TGGCGCGGCTCAAGAGATGGAACGCTATGCGATGCTTCCGCTTCCGCGCGCGCCCA 922  
QY 926 GCGTCAACGTGACCTCAAGTACATCAAGGCGAAGTCTTCCGGAAGGCGAGGCGG 985  
Db 923 GCGTCAACGTGACCTCAAGTACATCAAGGCGAAGTCTTCCGGAAGGCGAGGCGG 982  
QY 986 AGAGGGAACAAGTGCACATCCAGTGAAGGAGGAGTGTGAGGCGGCGCTTGGCCAGTGC 1045  
Db 983 AGAGGGAACAAGTGCACATCCAGTGAAGGAGGAGTGTGAGGCGGCGCTTGGCCAGTGC 1042  
QY 1046 TTCAAGGAGGTGGCCCAAGATGCCCATGTGTGCGATCTTCCCAACGAGGCGCGGAGCA 1105  
Db 1043 TTCAAGGAGGTGGCCCAAGATGCCCATGTGTGCGATCTTCCCAACGAGGCGCGGAGCA 1102  
QY 1106 GTCTTGTTCAAGACCTTAGGCAACGAGCTGAGGCGCGCGGCGCTTGGCCAGTGC 1165  
Db 1103 GTCTTGTTCAAGACCTTAGGCAACGAGCTGAGGCGCGCGGCGCTTGGCCAGTGC 1162  
QY 1166 TCGTCTGCTTCTCAAGACCTTCTGAGTCCGCTTGTCCAGAGCTCTTGGTGGTTTCA 1225  
Db 1163 TCGTCTGCTTCTCAAGACCTTCTGAGTCCGCTTGTCCAGAGCTCTTGGTGGTTTCA 1222  
QY 1226 CTCTCTGTGGAAGACATCTCTGACGCTCAAGAGTTAGGAGAGACTCAAGTTACA 1285  
Db 1223 CTCTCTGTGGAAGACATCTCTGACGCTCAAGAGTTAGGAGAGACTCAAGTTACA 1282  
QY 1286 CCTTCTCTCTGGGGTTGGAAGAAATGTTAGTGCAGAGGGGTGAGGATTGCTGCTCA 1345  
Db 1283 CCTTCTCTCTGGGGTTGGAAGAAATGTTAGTGCAGAGGGGTGAGGATTGCTGCTCA 1342  
QY 1346 TATGAGAGCTCTGGAACAGCTCAGAGATGAAAAGACAGAGAGGCGAGATGAGAA 1405  
Db 1343 TATGAGAGCTCTGGAACAGCTCAGAGATGAAAAGACAGAGAGGCGAGATGAGAA 1402  
QY 1406 GTCTCTCTCTCTGCGCATTAACCCAGCTTGGTTGGGTGCGAGCTGGAGAACTTCTC 1465  
Db 1403 GTCTCTCTCTCTGCGCATTAACCCAGCTTGGTTGGGTGCGAGCTGGAGAACTTCTC 1462  
QY 1466 TCCAGAGCTGCAACTTTAGCTCTGTTGAGCTGCTGCAACCCCTCCCAACCCCA 1525  
Db 1463 TCCAGAGCTGCAACTTTAGCTCTGTTGAGCTGCTGCAACCCCTCCCAACCCCA 1522  
QY 1526 GCAACAACAAGTTGAGCCCGAGCTGCGCTGACATGAGCCAGTGAAGTCTGTGTG 1585  
Db 1523 GCAACAACAAGTTGAGCCCGAGCTGCGCTGACATGAGCCAGTGAAGTCTGTGTG 1582  
QY 1586 AAGGGGCGTGGCCACACTTCTAGACACGCGCCACCACTTAGACACGCGCCACTCTG 1645  
Db 1583 AAGGGGCGTGGCCACACTTCTAGACACGCGCCACCACTTAGACACGCGCCACTCTG 1642

QY 1646 ACCGCTTCCCTCAGCTCTCTCTCTAGATGCTCCGCGCCGACAGTTGCTGTGTG 1705  
Db 1643 ACCGCTTCCCTCAGCTCTCTCTCTAGATGCTCCGCGCCGACAGTTGCTGTGTG 1702  
QY 1706 TTGAGCTGTTTTGTTGTCATGTATAGTAGTAAATGAAATCATTTGATCTGTAAAGC 1765  
Db 1703 TTGAGCTGTTTTGTTGTCATGTATAGTAGTAAATGAAATCATTTGATCTGTAAAGC 1762  
QY 1766 CTATGACTCCCTCTCTTGGCGAGGCTTCAACCCAGTTCAAGATCCAGCGCTCCACCCGG 1825  
Db 1763 CTATGACTCCCTCTCTTGGCGAGGCTTCAACCCAGTTCAAGATCCAGCGCTCCACCCGG 1822  
QY 1826 ACGCTTCTCTCTCTGCTCCCAACAGGTTTCCGCGGCTGTTTGGAGCTAGCATTTGA 1885  
Db 1823 ACGCTTCTCTCTCTGCTCCCAACAGGTTTCCGCGGCTGTTTGGAGCTAGCATTTGA 1882  
QY 1886 CCTCCGCAATTTAGCTTCCAGCGTTTACAGACATTTGACAAAGCGTGGGGTGGGAGCCA 1945  
Db 1883 CCTCCGCAATTTAGCTTCCAGCGTTTACAGACATTTGACAAAGCGTGGGGTGGGAGCCA 1942  
QY 1946 GGAATGCTTTTTTTTAAATGCTTCCATTTCAAGAGATACCAACGAGCTCGAGGGAGC 2005  
Db 1943 GGAATGCTTTTTTTTAAATGCTTCCATTTCAAGAGATACCAACGAGCTCGAGGGAGC 2002  
QY 2006 ACGATGAGACACAGGCGCCCACTTTGTCCCTAGCAATTCAGGGTACAGCTCCACTAG 2065  
Db 2003 ACGATGAGACACAGGCGCCCACTTTGTCCCTAGCAATTCAGGGTACAGCTCCACTAG 2062  
QY 2066 AACGAGGCTGCTCTTACTGTGCTGTTCTCAAGACATTTATTAAGACCTTACGGGTGC 2125  
Db 2063 AACGAGGCTGCTCTTACTGTGCTGTTCTCAAGACATTTATTAAGACCTTACGGGTGC 2122  
QY 2126 TGGGTTCATGTGTCTTAAGAAACCAAGAGGTCCCAAGTCTGCTGCGCGCGCCCT 2185  
Db 2123 TGGGTTCATGTGTCTTAAGAAACCAAGAGGTCCCAAGTCTGCTGCGCGCGCCCT 2182  
QY 2186 GCTGCCCAACCACTTCTGCAACACAGCGGTGGGAGGCGGGAGGAGACGCTGGAGC 2245  
Db 2183 GCTGCCCAACCACTTCTGCAACACAGCGGTGGGAGGCGGGAGGAGACGCTGGAGC 2242  
QY 2246 CAGAACTGAGCTGGAGGAGTCCGACAGAAAGCTCAGGCGGGGTCTTCTCTGTGTC 2305  
Db 2243 CAGAACTGAGCTGGAGGAGTCCGACAGAAAGCTCAGGCGGGGTCTTCTCTGTGTC 2302  
QY 2306 CGGATTTGGGTATGCTGGGTGACCAACATGTAAGTCAAGGATGCTGTGTAACCATTA 2365  
Db 2303 CGGATTTGGGTATGCTGGGTGACCAACATGTAAGTCAAGGATGCTGTGTAACCATTA 2362  
QY 2366 AACCAAGGCGCTTGTATCAGCTCTTAAACAAGTATTTTGTATTTTAACTCTTAAA 2425  
Db 2363 AACCAAGGCGCTTGTATCAGCTCTTAAACAAGTATTTTGTATTTTAACTCTTAAA 2422  
QY 2426 CATATGAAGTTTAAAGGCGCTTAAGAAACCTTAGTATCTTATTTGGGTCTTCTGAGG 2485  
Db 2423 CATATGAAGTTTAAAGGCGCTTAAGAAACCTTAGTATCTTATTTGGGTCTTCTGAGG 2482  
QY 2486 TTCAAGAGGGTATGTAATTTCTCCAGGTACACACAAAGTCTGTGGGTGGCAAGCA 2545  
Db 2483 TTCAAGAGGGTATGTAATTTCTCCAGGTACACACAAAGTCTGTGGGTGGCAAGCA 2542  
QY 2546 AGTATGAGCTGGGATTCAGTACATACAGATGTGCTCCCTCTTGAAGCTTGGCCCC 2605  
Db 2543 AGTATGAGCTGGGATTCAGTACATACAGATGTGCTCCCTCTTGAAGCTTGGCCCC 2602  
QY 2606 TGGGGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTTCTCCATGATCATGCTGTGA 2665  
Db 2603 TGGGGCTTCAAGGCTTTGGGACATCTTGTCTCAACCTTCTCCATGATCATGCTGTGA 2662  
QY 2666 GGTTCCTCTGATATTTGTGTACACATGCCCCATGTAATTAACAAGTACACACAGATGAC 2725  
Db 2663 GGTTCCTCTGATATTTGTGTACACATGCCCCATGTAATTAACAAGTACACACAGATGAC 2722  
QY 2726 ACAAGATGTACATGCTCCAGGCCAGCTCTGCAATGCTGCACTGCAACCCAGGCTT 2785

|||||  
Db 2723 ACAAGATGATACATCTCCAGCCCAAGCTCTGATACCTGACCTGCAACCCAGGCTT 2782  
|||  
Qy 2786 GGGCCCTGCTGGTCTGTGTCTAAAGACAGAGCTCCAAACCTGCTCTGTCCCTTTCC 2845  
|||  
Db 2783 GGGCCCTGCTGGTCTGTGTCTAAAGACAGAGCTCCAAACCTGCTCTGTCCCTTTCC 2842  
|||  
Qy 2846 CACCCACTGCTGAGCTTCTGAGCAGACAGAGTACCTGGCTGCAACGGTGTGGGCC 2905  
|||  
Db 2843 CACCCACTGCTGAGCTTCTGAGCAGACAGAGTACCTGGCTGCAACGGTGTGGGCC 2902  
|||  
Qy 2906 GCTCTACCCAGGACAGACCCCGGCAACATGATCTCCGTGATCACTATCAATAAAGT 2965  
|||  
Db 2903 GCTCTACCCAGGACAGACCCCGGCAACATGATCTCCGTGATCACTATCAATAAAGT 2962  
|||  
Qy 2966 GGTGTGTTACA 2976  
|||  
Db 2963 GGTGTGTTACA 2973  
|||  
RESULT 4  
AX393244 2832 bp DNA linear PAT 23-MAR-2002  
LOCUS AX393244  
DEFINITION Sequence 174 from Patent WO0210217.  
ACCESSION AX393244  
VERSION AX393244.1 GI:19701294  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS St Croix, B., Kinzler, K.W. and Vogelstein, B.  
TITLE Endothelial cell expression patterns  
JOURNAL Patent: WO 0210217-A 174 07-FEB-2002;  
The Johns Hopkins University (US)  
FEATURES  
source 1..2832  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 89.5%; Score 2757; DB 6; Length 2832;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 199 CAGCCCGAGCAGATGATGAAGACTTGTCCAGCGGAACTGCAAGTGTGCGGCC 258  
|||  
Db 25 CAGCCCGAGCAGATGATGAAGACTTGTCCAGCGGAACTGCAAGTGTGCGGCC 84  
|||  
Qy 259 AAAAATCATACCGCATGTGTGTCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGT 318  
|||  
Db 85 AAAAATCATACCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 144  
|||  
Qy 319 TCTGCTCTCTCAATGCGCGCTTTGAGGACCAATGACACCCCAATGAGGAACTTCCAC 378  
|||  
Db 145 TCTGCTCTCTCAATGCGCGCTTTGAGGACCAATGACACCCCAATGAGGAACTTCCAC 204  
|||  
Qy 379 CGTAAGGATATACATTCGCGGCGGAGATGATACAGCTGCAATCTGATATCTGCGC 438  
|||  
Db 205 CGTAAGGATATACATTCGCGGCGGAGATGATACAGCTGCAATCTGATATCTGCGC 264  
|||  
Qy 439 AACCAACCTCTCCCGCATGCGAGGCTGTCAATCTTCAAGGGAGTGTCTTATCTTG 498  
|||  
Db 265 AACCAACCTCTCCCGCATGCGAGGCTGTCAATCTTCAAGGGAGTGTCTTATCTTG 324  
|||  
Qy 499 GTGTTCAGCTGATTAACCGGAGTCTTGTGATGATGAGTCAAGCGCTTCAAGAACATC 558  
|||  
Db 325 GTGTTCAGCTGATTAACCGGAGTCTTGTGATGATGAGTCAAGCGCTTCAAGAACATC 384  
|||  
Qy 559 CTGAGGTCAAGTCTGCTGTAAGAACAAAGCAAGAGCGCGGAGTCCCATGATC 618  
|||

|||||  
Db 385 CTGAGGTCAAGTCTGCTGTAAGAAACAAGACCAAGAGCGCGAGTGTCCCATGTC 444  
|||  
Qy 619 ATCTGTGGCAACAAGAACCAACGCGAGCTGTGCGCCAGGTGTCCACACCGAGGCC 678  
|||  
Db 445 ATCTGTGGCAACAAGAACCAACGCGAGCTGTGCGCCAGGTGTCCACACCGAGGCC 504  
|||  
Qy 679 GAGCTGTGTGTGTGCGGCGAGAGAACTCGCTTACTTCAAGGTGTGTGCGCAAGAAC 738  
|||  
Db 505 GAGCTGTGTGTGTGCGGCGAGAGAACTGTGCGCTTACTTCAAGGTGTGTGCGCAAGAAC 564  
|||  
Qy 739 ACCAAGTGAAGAGATGTTCTAGTGTCTTGAAGATGAGCAAGCTGCAACAGATG 798  
|||  
Db 565 ACCAAGTGAAGAGATGTTCTAGTGTCTTGAAGATGAGCAAGCTGCAACAGATG 624  
|||  
Qy 799 AGCCCGGCTGTGATGCGAAGATCTCGTGTGATGAGTCAAGCTTCCACCCAGGCC 858  
|||  
Db 625 AGCCCGGCTGTGATGCGAAGATCTCGTGTGATGAGTCAAGCTTCCACCCAGGCC 684  
|||  
Qy 859 TTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGATGCTGTGCGCTTGTGCGCG 918  
|||  
Db 685 TTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGATGCTGTGCGCTTGTGCGCG 744  
|||  
Qy 919 CGCCCGAGCTCAACAGTGACTTCAAGTACATCAAGGCAAGGCTTGTGGAAGGCGAG 978  
|||  
Db 745 CGCCCGAGCTCAACAGTGACTTCAAGTACATCAAGGCAAGGCTTGTGGAAGGCGAG 804  
|||  
Qy 979 GCGCGTGAAGGAGCAAGTGAACATCAATGAGCGAGGATGTGTGGGGGGGGCTTGGC 1038  
|||  
Db 805 GCGCGTGAAGGAGCAAGTGAACATCAATGAGCGAGGATGTGTGGGGGGGGCTTGGC 864  
|||  
Qy 1039 CAGTGCCTTGAAGGAGGTGCGCCAGATGCGCACTGTGCGCATCTCCCAACGAGGCC 1098  
|||  
Db 865 CAGTGCCTTGAAGGAGGTGCGCCAGATGCGCACTGTGCGCATCTCCCAACGAGGCC 924  
|||  
Qy 1099 GCGAGCATGTTGTGCTTCAACAGCTTGAAGCAACAGTGAAGGCGCGCGCTGCTTC 1158  
|||  
Db 925 GCGAGCATGTTGTGCTTCAACAGCTTGAAGCAACAGTGAAGGCGCGCGCTGCTTC 984  
|||  
Qy 1159 CGCAGATGCTGTGCTTCAACAGCTTGAAGCAACAGTGAAGGCGCGCGCTGCTTC 1218  
|||  
Db 985 CGCAGATGCTGTGCTTCAACAGCTTGAAGCAACAGTGAAGGCGCGCGCTGCTTC 1044  
|||  
Qy 1219 GTTTCATCTCTCTGTGAGGAGACACATCTGTGAGGCTTCAAGGTTAGGACAGACTCA 1278  
|||  
Db 1045 GTTTCATCTCTCTGTGAGGAGACACATCTGTGAGGCTTCAAGGTTAGGACAGACTCA 1104  
|||  
Qy 1279 AGTTACACTTCTCTCTGTGAGGAGGTTGGAAGAAATGTTATGCCAGAGGAGATGTC 1338  
|||  
Db 1105 AGTTACACTTCTCTCTGTGAGGAGGTTGGAAGAAATGTTATGCCAGAGGAGATGTC 1164  
|||  
Qy 1339 TGGCTCATATGAGAGCTCTGTGAGCAAGCTCAGAGATGAAGAAAGGACACAGAGGCGAT 1398  
|||  
Db 1165 TGGCTCATATGAGAGCTCTGTGAGCAAGCTCAGAGATGAAGAAAGGACACAGAGGCGAT 1224  
|||  
Qy 1399 GAGAAAGTCTCTCTCTGTGAGATTAACCAAGCTGTGTTGGGTGGGAGAGTGGAGA 1458  
|||  
Db 1225 GAGAAAGTCTCTCTCTGTGAGATTAACCAAGCTGTGTTGGGTGGGAGAGTGGAGA 1284  
|||  
Qy 1459 ACTTCTCTCCAGCGCTCAACTTCAAGCTGTGTTCAAGTGTGCAAGCTTGTGCA 1518  
|||  
Db 1285 ACTTCTCTCCAGCGCTCAAGCTTCAAGCTGTGTTCAAGTGTGCAAGCTTGTGCA 1344  
|||  
Qy 1519 ACCCCAGCAACACAGATGAGGCGCGAGCTGTGAGTGAATGAGCTTGAAGCTTGT 1578  
|||  
Db 1345 ACCCCAGCAACACAGATGAGGCGCGAGCTGTGAGTGAATGAGCTTGAAGCTTGT 1404  
|||  
Qy 1579 GTGTGTAAGGGGGCGGAGCACAACCTCTAGACCAAGGCCACCACTTGAAGCAAGGCCA 1638  
|||  
Db 1405 GTGTGTAAGGGGGCGGAGCACAACCTCTAGACCAAGGCCACCACTTGAAGCAAGGCCA 1464  
|||  
Qy 1639 CTTCCTGACCGGCTTCTGAGCTCTCTCTAGTCTCTCTGCGCGGACAGTGTGCTTT 1698  
|||  
Db 1465 CTTCCTGACCGGCTTCTGAGCTCTCTCTAGTCTCTCTGCGCGGACAGTGTGCTTT 1524  
|||



QY	1639	GTGTGGTGCAGCGCTTTTGCTGTGCATGATAGATAGTAAAGAAATGAATCATTTGACTGC	1758
Db	1525	GTTTGCTGTGCAGCGCTTTTGCTGTGCATGATAGATAGTAAAGAAATCATTTGACTGC	1584
QY	1759	TAAAGCCTAGTACTCCCTCTTGAGCCAGGCGCTTACCAGTTCCAGTCCAGCGGCTCC	1818
Db	1585	TAAAGCCTAGTACTCCCTCTTGAGCCAGGCGCTTACCAGTTCCAGTCCAGCGGCTCC	1644
QY	1819	ACCGGGAGCGGCTTCTCTCTCTGCTGCATCAACAGGGTTTCCGTGGCCGTTTTGACGCTAG	1878
Db	1645	ACCGGGAGCGGCTTCTCTCTCTGCTGCATCAACAGGGTTTCCGTGGCCGTTTTGACGCTAG	1704
QY	1879	ACATTGACCTCCGCCATTGAGCTCCACGGTTTACAGAAATTTGACACAGCTGGGGTGGG	1938
Db	1705	ACATTGACCTCCGCCATTGAGCTCCACGGTTTACAGAAATTTGACACAGCTGGGGTGGG	1766
QY	1939	CAGGCTCAGAGTCTGCTTTTTTTTAATGCTCCATTTCACAGAGATACACCGAGCTCGG	1998
Db	1765	CAGGCTCAGAGTCTGCTTTTTTTTAATGCTCCATTTCACAGAGATACACCGAGCTCGG	1824
QY	1999	AGGGGACACGATGAGACACAGGCGCCACCTTTGCTCCCTAGCAATTAAGGGTTCAGCTC	2058
Db	1825	AGGGGACACGATGAGACACAGGCGCCACCTTTGCTCCCTAGCAATTAAGGGTTCAGCTC	1884
QY	2059	CACCTAGAACCAAGGCTGCCCTTCTACTGTGCTGTTCCTCAAGCATTTTATTAAGCACTTAC	2118
Db	1885	CACCTAGAACCAAGGCTGCCCTTCTACTGTGCTGTTCCTCAAGCATTTTATTAAGCACTTAC	1944
QY	2119	TGGGTGCTGGGGTTTCACTGTGTCTTAGAGAAACAAAGAGGTCCCAAGTCTGGGCTCTGGC	2178
Db	1945	TGGGTGCTGGGGTTTCACTGTGTCTTAGAGAAACAAAGAGGTCCCAAGTCTGGGCTCTGGC	2004
QY	2179	CGCCCTGCTGCTCCCAACCACTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGCAGC	2238
Db	2005	CGCCCTGCTGCTCCCAACCACTTCTGACACACAGCGGTGGGAGGCGGGAGAGAGCAGC	2064
QY	2239	TGGGATCCCAAGAACTGAGCGCTGGAGAGGATTCGACAGAAAGCTCAGGGCGGGCTTCTTCC	2298
Db	2065	TGGGATCCCAAGAACTGAGCGCTGGAGAGGATTCGACAGAAAGCTCAGGGCGGGCTTCTTCC	2124
QY	2299	TTGTGTCGCGGAGTTGGGCTATGCTGCGGGATCACACATGTACTCAGGCATGCGGGGTTTTGA	2358
Db	2125	TTGTGTCGCGGAGTTGGGCTATGCTGCGGGATCACACATGTACTCAGGCATGCGGGGTTTTGA	2184
QY	2359	ACCCTAATAACCAAGGCCCTTGTGCATCGCTCTTAAACAGATATATTTTGTATTTTAAATCT	2418
Db	2185	ACCCTAATAACCAAGGCCCTTGTGCATCGCTCTTAAACAGATATATTTTGTATTTTAAATCT	2244
QY	2419	CTCTAATACTATTTAAAGTTTAAAGGCCCTTAAGAACTTATGTATTTTCTATTTGGGCTT	2478
Db	2245	CTCTAATACTATTTAAAGTTTAAAGGCCCTTAAGAACTTATGTATTTTCTATTTGGGCTT	2304
QY	2479	TCTGAGGTTTCAGAGGGGTAAAGTAACTTCTCCAGGTTCACACAGAACTGTGGGTGGC	2538
Db	2305	TCTGAGGTTTCAGAGGGGTAAAGTAACTTCTCCAGGTTCACACAGAACTGTGGGTGGC	2364
QY	2539	AGAAAGCAAGCTAGCGGCTGGGCACTCAGTACATACACAGATGTGCTCCCTCTCTTGATGCT	2598
Db	2365	AGAAAGCAAGCTAGCGGCTGGGCACTCAGTACATACACAGATGTGCTCCCTCTCTTGATGCT	2424
QY	2599	TGGGCCCTGGGGGCTTTCAGGGCTTTTGGGACATCTTGTGCTCAAACTCTCCCTAGATCAG	2658
Db	2425	TGGGCCCTGGGGGCTTTCAGGGCTTTTGGGACATCTTGTGCTCAAACTCTCCCTAGATCAG	2484
QY	2659	TCTGTGAGGGTCCCTGTAGATATTTGTGTACACATGCGCCATGTATATACAGTACACACA	2718
Db	2485	TCTGTGAGGGTCCCTGTAGATATTTGTGTACACATGCGCCATGTATATACAGTACACACA	2544
QY	2719	GATGTACACACAGATGTACACATGCTCAGGCCACCGAGCTCTGACATCTGTGACCTGTGACCC	2778
Db	2545	GATGTACACACAGATGTACACATGCTCAGGCCACCGAGCTCTGACATCTGTGACCTGTGACCC	2604

Qy	2779	CAGCCTTGAGCCCTGCTCGCGCTGTGTGCTCAAGAGAGAGAGCTCCAAACCTGCTGTGTC	2833		
Db	2605	CAGCCTTGAGCCCTGCTCGCGCTGTGTGCTCAAGAGAGAGAGCTCCAAACCTGCTGTGTC	2666		
Qy	2839	CCTTCCCAACCACTGCTGAGCCTTCTGAGCAGACCAAGTACCTTGGCTGCACCGGTGT	2898		
Db	2665	CCTTCCCAACCACTGCTGAGCCTTCTGAGCAGACCAAGTACCTTGGCTGCACCGGTGT	2722		
Qy	2899	GTGGCCGCTCTCACCCAGGACAGCCCGGCACCATGTGATCTCGGTATACATATCAAT	2955		
Db	2725	GTGGCCGCTCTCACCCAGGACAGCCCGGCACCATGTGATCTCGGTATACATATCAAT	2788		
Qy	2959	AAAAGTGCGTTGTTCACAAAAA	3006		
Db	2785	AAAAGTGCGTTGTTCACAAAAA	2832		
RESULT 5	AF279143	2832 bp	mRNA	linear	PRI 23-AUG-2000
LOCUS	AF279143				
DEFINITION	Homo sapiens tumor endothelial marker 2 (TEM2) mRNA, complete cds.				
ACCESSION	AF279143				
VERSION	AF279143.1				
KEYWORDS	GI:9857401				
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 2832) St. Croix,B., Rago,C., Velculescu,V., Traverso,G., Romans,K.E., Montgomerie,E., Lal,A., Riggin,G.J., Lengauer,C., Vogelstein,B. and Kinzler,K.W.				
TITLE	Genes expressed in human tumor endothelium				
JOURNAL	Science 289 (5482), 1197-1202 (2000)				
MEDLINE	20407466				
PubMed	10947988				
REFERENCE	2 (bases 1 to 2832) St. Croix,B., Vogelstein,B. and Kinzler,K.W.				
AUTHORS	Direct Submission				
TITLE	Submitted (15-JUN-2000) Johns Hopkins Oncology Center, Johns Hopkins University, 1650 Orleans Street, Baltimore, MD 21231, USA				
JOURNAL	Location/Qualifiers				
FEATURES	1..2832				
source	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="2"				
	/map="22q13.1"				
gene	1..2832				
	/gene="TEM2"				
CDS	1..837				
	/gene="TEM2"				
	/note="TEM2 protein"				
	/codon_start=1				
	/product="tumor endothelial marker 2"				
	/protein_id="AA000688.1"				
	/db_xref="GI:9857402"				
	/translation="MPASLLDLPAMMKTLISGNCITLSPAKNSYRMVTVIGASRVGKSSIVSRFLNGRFEDQYPTLTEDFHRKYNIKGMVQDILDTSGNHFPFPMRRLSILTTGADVIFRSLDNRESFDEVRQLQKILIEVSCIKNKTKEAAELPMVTCGNKHGELCROYPTTEAEELVSGDENCAFEVSAKNTVDSEFVYLFMAKLPHMSPALRKISVQYGDAPFHRPFCRRVRVEMDAYGVGVSPFARPSVNSDKITIKAKVLRBEGARERDKCTIQ"				
ORIGIN					
Query Match	89.5%;	Score 2757;	DB 9;	Length 2832;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2807;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
Qy	199	CAGCCCCGAGCCCATGATGAAGACTTGTCCAGCGGGAACTGCACGCTCACTGTGCCCCGCC	258		
Db	25	CAGCCCCGAGCCCATGATGAAGACTTGTCCAGCGGGAACTGCACGCTCACTGTGCCCCGCC	84		



QY 259 AAAAATCATACCGCATGAGTGTGCTGGAGTGCCTCTCGGTGGGCAAGAGCTCCATCGT 318  
Db 85 AAAAATCATACCGCATGAGTGTGCTGGAGTGCCTCTCGGTGGGCAAGAGCTCCATCGT 144  
QY 319 TCTGCTTCTCTCAATGGCGCTTTGAGGACCAAGTACACACCCCAATCGAGACTTTCAC 378  
Db 145 TCTGCTTCTCTCAATGGCGCTTTGAGGACCAAGTACACACCCCAATCGAGACTTTCAC 204  
QY 379 CGTAAGGTATACAAATCCGCGCGCATATGATCAAGTGTGACATCCCTGGATACCTCGGC 438  
Db 205 CGTAAGGTATACAAATCCGCGCGCATATGATCAAGTGTGACATCCCTGGATACCTCGGC 264  
QY 439 AACCAACCCCTTCCCGCATGCGCAAGAGTGTCCATCTTCAAGAGGATGCTTATCTG 498  
Db 265 AACCAACCCCTTCCCGCATGCGCAAGAGTGTCCATCTTCAAGAGGATGCTTATCTG 324  
QY 499 GTGTTCAGCTGTGATTAACCGGAGTCCCTTGATGAGGTCAAGCGCTTTCAGAAAGCATATC 558  
Db 325 GTGTTCAGCTGTGATTAACCGGAGTCCCTTGATGAGGTCAAGCGCTTTCAGAAAGCATATC 384  
QY 559 CTGGAAGTCAAGTCTCTGCTGAGAAACAAGACCAAGAGCGCGGAGCTGCCCATGCTC 618  
Db 385 CTGGAAGTCAAGTCTCTGCTGAGAAACAAGACCAAGAGCGCGGAGCTGCCCATGCTC 444  
QY 619 ATCTGTGGCAACAAAGACGACACGCGAGCTGTGCCAGAGTGCACACCAAGAGGCC 678  
Db 445 ATCTGTGGCAACAAAGACGACACGCGAGCTGTGCCAGAGTGCACACCAAGAGGCC 504  
QY 679 GAGCTGTGTGTGCGGCGACGAGAACTCGCTACTTTCAGAGTGTGCGCCAAAGAAAC 738  
Db 505 GAGCTGTGTGTGCGGCGACGAGAACTCGCTACTTTCAGAGTGTGCGCCAAAGAAAC 564  
QY 739 ACCAAGTGAACGAGATGTTCTACGTGCTTTCAGCATGAGCCAAAGTGTCCACAGAGATG 798  
Db 565 ACCAAGTGAACGAGATGTTCTACGTGCTTTCAGCATGAGCCAAAGTGTCCACAGAGATG 624  
QY 799 AGCCCGCGCTGATGCGAAGATCTCCGTGATGAGTGAAGCGCTTTCACCCCAAGGCC 858  
Db 625 AGCCCGCGCTGATGCGAAGATCTCCGTGATGAGTGAAGCGCTTTCACCCCAAGGCC 684  
QY 859 TTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGAGCATGATCTCGCCCTTGCCCGC 918  
Db 685 TTCTGATGCGCGCGCTCAAGAGATGAGCGCTTATGAGCATGATCTCGCCCTTGCCCGC 744  
QY 919 CGCCCGAGCTCAACAGTGAATCTTCAAGTCAACAGCCAAAGTGTCTTTCGAGAGGCCAG 978  
Db 745 CGCCCGAGCTCAACAGTGAATCTTCAAGTCAACAGCCAAAGTGTCTTTCGAGAGGCCAG 804  
QY 979 GCCCGTAGAGGAGCAAGTGCACATCAAGTGAAGCGAGATGCTGGGGCGGGCTTGGC 1038  
Db 805 GCCCGTAGAGGAGCAAGTGCACATCAAGTGAAGCGAGATGCTGGGGCGGGCTTGGC 864  
QY 1039 CAGTGCCTTGAAGGAGTGGCCCAAGTGCACATCTGTGCGCATCTCCCAAGAGGCC 1098  
Db 865 CAGTGCCTTGAAGGAGTGGCCCAAGTGCACATCTGTGCGCATCTCCCAAGAGGCC 924  
QY 1099 GCGAGCAGTCTTGTTCACACACTTGAAGCACCACTGAGAGGCCCGGGGCTGCGCTC 1158  
Db 925 GCGAGCAGTCTTGTTCACACACTTGAAGCACCACTGAGAGGCCCGGGGCTGCGCTC 984  
QY 1159 CGCAGATTCGTCGCTTCTCAAGCTTTCAGAGCTTTCGATCGGCTTTCACAGCTCTTGGTG 1218  
Db 985 CGCAGATTCGTCGCTTCTCAAGCTTTCAGAGCTTTCGATCGGCTTTCACAGCTCTTGGTG 1044  
QY 1219 GTTTCATCTCTCTGTGGAGAGACACATCTTTCAGAGCTTTCAGAGTATAGGAGAGACTCA 1278  
Db 1045 GTTTCATCTCTCTGTGGAGAGACACATCTTTCAGAGCTTTCAGAGTATAGGAGAGACTCA 1104  
QY 1279 AGTTAACAATCTCTCTCTGTGGAGTGAAGAAATGTTGATGCCAGAGGGGTGAAGATTTGC 1338  
Db 1105 AGTTAACAATCTCTCTCTGTGGAGTGAAGAAATGTTGATGCCAGAGGGGTGAAGATTTGC 1164  
QY 1339 TGCGTATATGAGACCTCTCTGAGCAAGCTCAAGATGAAAAGACACAGAAAGCCAGAT 1398

Db 1165 TGCGTATATGAGACCTCTCTGAGCAAGCTCAAGATGAAAAGACACAGAAAGCCAGAT 1224  
QY 1339 GAGAAAGTCTCTCTCTCTGTGGATTAACACCCAGCTTGGTTGGGTGGAGCTGGGAGA 1458  
Db 1225 GAGAAAGTCTCTCTCTCTGTGGATTAACACCCAGCTTGGTTGGGTGGAGCTGGGAGA 1284  
QY 1459 ACTTCTCTCCAGCCCTGCAACTTACGCTGTGTTACGCTCTTTCAGTCCCTTCAGCCCTCC 1518  
Db 1285 ACTTCTCTCCAGCCCTGCAACTTACGCTGTGTTACGCTCTTTCAGTCCCTTCAGCCCTCC 1344  
QY 1519 ACCCCAGACACACAAAGTTGGCCCAAGCTGCGCTGACATTTAGACAGTGAATCTT 1578  
Db 1345 ACCCCAGACACACAAAGTTGGCCCAAGCTGCGCTGACATTTAGACAGTGAATCTT 1404  
QY 1579 GTGTCTGAAGGGGGCGTGGGACACCTCTCTAGACACAGCCACACCTTAAACACCGCCA 1638  
Db 1405 GTGTCTGAAGGGGGCGTGGGACACCTCTCTAGACACAGCCACACCTTAAACACCGCCA 1464  
QY 1639 CTTCTGACCGGCTCTGAGCTCTCTCTAGTCCCTCGCGCCGACAGTTGCTTT 1698  
Db 1465 CTTCTGACCGGCTCTGAGCTCTCTCTAGTCCCTCGCGCCGACAGTTGCTTT 1524  
QY 1699 GTTGTGTTGACAGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758  
Db 1525 GTTGTGTTGACAGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584  
QY 1759 TAAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818  
Db 1585 TAAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1644  
QY 1819 ACCCGGAGCGCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1878  
Db 1645 ACCCGGAGCGCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704  
QY 1879 ACAATTGACTCGCGCTTGAAGTCAAGGTTTACAGAAATGACAAAGCGTGGGTGGG 1938  
Db 1705 ACAATTGACTCGCGCTTGAAGTCAAGGTTTACAGAAATGACAAAGCGTGGGTGGG 1764  
QY 1939 CAGGCCAGAGCTGTTTTTTTAAATGATGCCATTTTCAAGAGATTAACACAGAGACTCGG 1998  
Db 1765 CAGGCCAGAGCTGTTTTTTTAAATGATGCCATTTTCAAGAGATTAACACAGAGACTCGG 1824  
QY 1999 AGGGGACAGATGAGACCAAGGCCCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2058  
Db 1825 AGGGGACAGATGAGACCAAGGCCCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1884  
QY 2059 CACCTAGAACCAAGCTGCTCTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2118  
Db 1885 CACCTAGAACCAAGCTGCTCTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944  
QY 2119 TGGGTGCTGGGTTCACTGTGCTCTAGAGAACCAAGAGGTTCCCAAGTCTGAGCTTGGCC 2178  
Db 1945 TGGGTGCTGGGTTCACTGTGCTCTAGAGAACCAAGAGGTTCCCAAGTCTGAGCTTGGCC 2004  
QY 2179 CGCCCTGCTGCGCCCAACACTTCTGACACAGCGGTGGGAGGCGGGGAGGAGGAGC 2238  
Db 2005 CGCCCTGCTGCGCCCAACACTTCTGACACAGCGGTGGGAGGCGGGGAGGAGGAGC 2064  
QY 2239 TGGGACCCAGAACTGAGCTGAGGAGGATCGACAGAAAGCTCAAGGCGGGCTTCTCC 2298  
Db 2065 TGGGACCCAGAACTGAGCTGAGGAGGATCGACAGAAAGCTCAAGGCGGGCTTCTCC 2124  
QY 2299 TTGTGCTCGGGATTTGGGCTATGCTGGGTACACATGATCTCAGGACATGAGTGGTTTGA 2358  
Db 2125 TTGTGCTCGGGATTTGGGCTATGCTGGGTACACATGATCTCAGGACATGAGTGGTTTGA 2184  
QY 2359 ACCCATTAACCAAGGCGCTTGCATGAGCTTAAACAAGTATTTTATTTAATCT 2418  
Db 2185 ACCCATTAACCAAGGCGCTTGCATGAGCTTAAACAAGTATTTTATTTAATCT 2244  
QY 2419 CTTTAAACATATGAACTTTTAAAGGCTTAAAGAACCTTAAATCTTCTTAAATTTGGGCTT 2478

Db 2245 CTCTAATATATTAAGTTTATAGGGCCCTAAGAACTTAGATCTTTATATGGGCTT 2304  
Qy 2479 TCTGAGTTTCAGAGAGGGTAACTTCTCTCAGAGTCAACAGAACTCTGTGGGTGC 2538  
Db 2305 TCTGAGTTTCAGAGAGGGTAACTTCTCTCAGAGTCAACAGAACTCTGTGGGTGC 2364  
Qy 2539 AGAAGCAAGCTAGAGAGGGTAACTTCTCTCAGAGTCAACAGAACTCTGTGGGTGC 2598  
Db 2355 AGAAGCAAGCTAGAGAGGGTAACTTCTCTCAGAGTCAACAGAACTCTGTGGGTGC 2424  
Qy 2599 TGGCCCTGGGGGCTTTCAGAGGCTTTGGGACATCTTGTCTCAACCTCTCTCAGATCAG 2658  
Db 2425 TGGCCCTGGGGGCTTTCAGAGGCTTTGGGACATCTTGTCTCAACCTCTCTCAGATCAG 2484  
Qy 2659 TCTGTAGAGGTTCCCTGTAGATATTTGTATACATATGCCATGTATATCAAGTACACA 2718  
Db 2485 TCTGTAGAGGTTCCCTGTAGATATTTGTATACATATGCCATGTATATCAAGTACACA 2544  
Qy 2719 GATGTACACAGATATGTATACATGTCTCCAGCCCTCTGATCTGATCTGATCTGATCTG 2778  
Db 2545 GATGTACACAGATATGTATACATGTCTCCAGCCCTCTGATCTGATCTGATCTGATCTG 2604  
Qy 2779 CAGCTTTGGCCCTCTGCTGCTGTGTCTCAAGCAGCAGCTCCAACTCTGCTCTGCTC 2838  
Db 2605 CAGCTTTGGCCCTCTGCTGCTGTGTCTCAAGCAGCAGCTCCAACTCTGCTCTGCTC 2664  
Qy 2839 CTTTCCCAACCCATGCTGCTGAGCTTCTGAGCAGACAGTACCTTGGCTGCACTGGTGT 2898  
Db 2665 CTTTCCCAACCCATGCTGCTGAGCTTCTGAGCAGACAGTACCTTGGCTGCACTGGTGT 2724  
Qy 2899 GTGCCCCGCTCTCAACCCAGCAGACCCGCGCACCATGATCTCCGTGTACATCTCAAT 2958  
Db 2725 GTGCCCCGCTCTCAACCCAGCAGACCCGCGCACCATGATCTCCGTGTACATCTCAAT 2784  
Qy 2959 AAAAGTGGGTTTGTATCAAAAAA 3006  
Db 2785 AAAAGTGGGTTTGTATCAAAAAA 2832

RESULT 6  
AX721069 3427 bp mRNA linear PAT 07-MAY-2003  
LOCUS  
DEFINITION Sequence 29 from Patent WO0220754.  
ACCESSION AX721069  
VERSION AX721069.1 GI:30421905  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1  
Lincoln, S.E., Altue, C.M., Dufour, G.E., Chalup, M.S., Hillman, J.L.,  
Jones, A., Yu, J.Y., Wright, R.U., Gietzen, D., Liu, Tommy, F., Yap, P.,  
Dahl, C.R., Momiya, M., Bradley, D., Rohatgi, S., Harris, B.,  
Roseberry, A.M., Gerstin, E.H., Peralta, C.H., David, M., Panzer, S.,  
Flores, V., Dafo, A., Marwaha, R., Chen, A., Chang, S.C. and Imman, R.R.  
Molecules for diagnosis and therapeutics  
Patent: WO 0220754-A 29 14-MAR-2002;  
JOURNAL Incyte Genomics, Inc. (US)  
TITLE Location/Qualifiers  
FEATURES  
source 1..3427  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: LG:337358.1:2000SEP08"

ORIGIN  
Query Match 83.4%; Score 2571; DB 6; Length 3427;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2961; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
Qy 26 GCCGCGCCAGCCCGGCTCCAGCAGCGCAGGAGGATCCCGCGCAGTGAACCGGG 85  
|||||

Db 459 GCCGCGCCAGCCCGGCTCCCGAGCAGCGCAGGAGAGATCCCGCGCAGTGAACCGGG 518  
Qy 86 AGCCACCAACAGCTCTGAGAGGCTCGCGGCTGAGAGCAGGAGCTCCCGCAGCTCC 145  
Db 519 AGCCACCAACAGCTCTGAGAGGCTCGCGGCTGAGAGCAGGAGCTCCCGCAGCTCC 578  
Qy 146 CGGGGCTTCAGGAGAGCTCTGAGAGCGGTGCACAGAGCGGGCGCGCATTCACAGCCCC 205  
Db 579 CGGGGCTTCAGGAGAGCTCTGAGAGCGGTGCACAGAGCGGGCGCGCATTCACAGCCCC 638  
Qy 206 GAGCCATGATGAGAGCTTTGTCCAGCGGAACTGACGCTCAGTGTGCCCGCAAAACT 265  
Db 639 GAGCCATGATGAGAGCTTTGTCCAGCGGAACTGACGCTCAGTGTGCCCGCAAAACT 698  
Qy 266 CATACCGATGT 325  
Db 699 CATACCGATGT 758  
Qy 326 TCCCTCAATGCGCGCTTTGAGAGCAGTATACACACCATGAGAGACTTCCACCGTAAG 385  
Db 759 TCCCTCAATGCGCGCTTTGAGAGCAGTATACACACCATGAGAGACTTCCACCGTAAG 818  
Qy 386 TATACCAATCCGCGGAGACATGTACAGCTCGACATCTGATATCTTGAACCAAC 445  
Db 819 TATACCAATCCGCGGAGACATGTACAGCTCGACATCTGATATCTTGAACCAAC 878  
Qy 446 CTTTCCCGCATGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 505  
Db 879 CTTTCCCGCATGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 938  
Qy 506 GCTGTGTAACCGGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565  
Db 939 GCTGTGTAACCGGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998  
Qy 566 TCAAGTCTGT 625  
Db 999 TCAAGTCTGT 1058  
Qy 626 GCACCAAGAACAGCAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 685  
Db 1059 GCACCAAGAACAGCAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1118  
Qy 686 TGT 745  
Db 1119 TGT 1178  
Qy 746 TGAAGAGATGTCTACGTCTCTTTCAGATGAGCAAGGCTGCAACAGATGAGCCCG 805  
Db 1179 TGAAGAGATGTCTACGTCTCTTTCAGATGAGCAAGGCTGCAACAGATGAGCCCG 1238  
Qy 806 CCGTCATGCGCAAGATCTCCGT 865  
Db 1239 CCGTCATGCGCAAGATCTCCGT 1298  
Qy 866 TGCCCGCCGCTCAAGAGATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 925  
Db 1299 TGCCCGCCGCTCAAGAGATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1358  
Qy 926 GCGTCAACAGTGAATCTCAAGTATCAAGGCTCTTTCAGAGAGGCTGAGGCTGCTG 985  
Db 1359 GCGTCAACAGTGAATCTCAAGTATCAAGGCTCTTTCAGAGAGGCTGAGGCTGCTG 1418  
Qy 986 AGAGGAGCAAGTGAATCTCAAGTATCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1045  
Db 1419 AGAGGAGCAAGTGAATCTCAAGTATCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1478  
Qy 1046 TTCAGGAGGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1105  
Db 1479 TTCAGGAGGTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1538  
Qy 1106 GTCTTTGTTCACAGCTTTAGGACACAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1165  
Db 1539 GTCTTTGTTCACAGCTTTAGGACACAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1598



## ORIGIN

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"Query Match 79.5%; Score 2449; DB 6; Length 11221;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 479 CAGGGAGTCTTCATCTGTGTTCAGCTGATTAACCGGAGTCTTGGATAGATCA 538  
Db CAGGGAGTCTTCATCTGTGTTCAGCTGATTAACCGGAGTCTTGGATAGATCA 538  
QY 7691 CAGGGAGTCTTCATCTGTGTTCAGCTGATTAACCGGAGTCTTGGATAGATCA 7750  
Db CAGGGAGTCTTCATCTGTGTTCAGCTGATTAACCGGAGTCTTGGATAGATCA 7750  
QY 539 AGGCGCTTCAAGACAGATCTGTGAGTCAAGTCTGCTGAAGAACAGCAAGAGAG 598  
Db AGGCGCTTCAAGACAGATCTGTGAGTCAAGTCTGCTGAAGAACAGCAAGAGAG 598  
QY 7751 AGGCGCTTCAAGACAGATCTGTGAGTCAAGTCTGCTGAAGAACAGCAAGAGAG 7810  
Db AGGCGCTTCAAGACAGATCTGTGAGTCAAGTCTGCTGAAGAACAGCAAGAGAG 7810  
QY 599 CGGCGAGTCTGCGCATCTGTGTGCAACAAGAACAGCAAGAGAGTGTGCGCC 658  
Db CGGCGAGTCTGCGCATCTGTGTGCAACAAGAACAGCAAGAGAGTGTGCGCC 658  
QY 7811 CGGCGAGTCTGCGCATCTGTGTGCAACAAGAACAGCAAGAGAGTGTGCGCC 7870  
Db CGGCGAGTCTGCGCATCTGTGTGCAACAAGAACAGCAAGAGAGTGTGCGCC 7870  
QY 659 AGGTGCCACCAAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 718  
Db AGGTGCCACCAAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 718  
QY 7871 AGGTGCCACCAAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7930  
Db AGGTGCCACCAAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7930  
QY 719 AGGTGTGGGCAAGAGAACCAACAGTGAAGATGTGTGTGTGTGTGTGTGTGTGT 778  
Db AGGTGTGGGCAAGAGAACCAACAGTGAAGATGTGTGTGTGTGTGTGTGTGTGT 778  
QY 7931 AGGTGTGGGCAAGAGAACCAACAGTGAAGATGTGTGTGTGTGTGTGTGTGTGT 7990  
Db AGGTGTGGGCAAGAGAACCAACAGTGAAGATGTGTGTGTGTGTGTGTGTGTGT 7990  
QY 779 CCAAGTGGCAACAGAGATGAGCCCGCTGATTCGCAAGATCTCGTGCAGTACGAGT 838  
Db CCAAGTGGCAACAGAGATGAGCCCGCTGATTCGCAAGATCTCGTGCAGTACGAGT 838  
QY 7991 CCAAGTGGCAACAGAGATGAGCCCGCTGATTCGCAAGATCTCGTGCAGTACGAGT 8050  
Db CCAAGTGGCAACAGAGATGAGCCCGCTGATTCGCAAGATCTCGTGCAGTACGAGT 8050  
QY 839 ACAGCTTCCACCCAGAGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 898  
Db ACAGCTTCCACCCAGAGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 898  
QY 8051 ACAGCTTCCACCCAGAGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8110  
Db ACAGCTTCCACCCAGAGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8110  
QY 899 TGTGCTGGCCCTTGTGCG 958  
Db TGTGCTGGCCCTTGTGCG 958  
QY 8111 TGTGCTGGCCCTTGTGCG 8170  
Db TGTGCTGGCCCTTGTGCG 8170  
QY 959 AGGTCTTTCGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1018  
Db AGGTCTTTCGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1018  
QY 8171 AGGTCTTTCGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8230  
Db AGGTCTTTCGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8230  
QY 1019 ATGCTGGGCG 1078  
Db ATGCTGGGCG 1078  
QY 8231 ATGCTGGGCG 8290  
Db ATGCTGGGCG 8290  
QY 1079 CATCTCCCAACAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1138  
Db CATCTCCCAACAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1138  
QY 8291 CATCTCCCAACAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8350  
Db CATCTCCCAACAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8350  
QY 1139 GGGCG 1198  
Db GGGCG 1198  
QY 8351 GGGCG 8410  
Db GGGCG 8410  
QY 1199 TTTGTCCACAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1258  
Db TTTGTCCACAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1258  
QY 8411 TTTGTCCACAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8470  
Db TTTGTCCACAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8470  
QY 1259 AAGAGTTAGGAGAGACTCAAGTTACACTTCTCTCTGCGGGTTGGAAGAAATGTTAT 1318  
Db AAGAGTTAGGAGAGACTCAAGTTACACTTCTCTCTCTGCGGGTTGGAAGAAATGTTAT 1318  
QY 8471 AAGAGTTAGGAGAGACTCAAGTTACACTTCTCTCTCTGCGGGTTGGAAGAAATGTTAT 8530  
Db AAGAGTTAGGAGAGACTCAAGTTACACTTCTCTCTCTGCGGGTTGGAAGAAATGTTAT 8530  
QY 1319 GCCAGAGGGGTGAGAGATTGCTGCGTCAATGAGAGCTCTCTGAGACAGCTTGAAGTAA 1378  
Db GCCAGAGGGGTGAGAGATTGCTGCGTCAATGAGAGCTCTCTGAGACAGCTTGAAGTAA 1378  
QY 8531 GCCAGAGGGGTGAGAGATTGCTGCGTCAATGAGAGCTCTCTGAGACAGCTTGAAGTAA 8590  
Db GCCAGAGGGGTGAGAGATTGCTGCGTCAATGAGAGCTCTCTGAGACAGCTTGAAGTAA 8590  
QY 1379 AAGAGACAGAGAGGAG 1438  
Db AAGAGACAGAGAGGAG 1438  
QY 8591 AAGAGACAGAGAGGAG 8650  
Db AAGAGACAGAGAGGAG 8650  
QY 1439 TTTGGGTGGCAGCTGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGCTGTGTTGAG 1498  
Db TTTGGGTGGCAGCTGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGCTGTGTTGAG 1498

Db TTTGGGTGGCAGCTGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGCTGTGTTGAG 8710  
QY 1499 CTGCTCTGCAACCCCTCTCCACCCCGAGCACAACACAGTGTGGCCCGCAGCTGCGCTG 1558  
Db CTGCTCTGCAACCCCTCTCCACCCCGAGCACAACACAGTGTGGCCCGCAGCTGCGCTG 8770  
QY 8711 CTGCTCTGCAACCCCTCTCCACCCCGAGCACAACACAGTGTGGCCCGCAGCTGCGCTG 8770  
Db CTGCTCTGCAACCCCTCTCCACCCCGAGCACAACACAGTGTGGCCCGCAGCTGCGCTG 8770  
QY 1559 ACAATTGAGCAGTGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1618  
Db ACAATTGAGCAGTGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1618  
QY 8771 ACAATTGAGCAGTGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8830  
Db ACAATTGAGCAGTGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8830  
QY 1619 CACCACTTGAACCAAGCCCACTCTGACCGCGTTCCTGAGCTCTCTCTGAGTCTCT 1678  
Db CACCACTTGAACCAAGCCCACTCTGACCGCGTTCCTGAGCTCTCTCTGAGTCTCT 1678  
QY 8831 CACCACTTGAACCAAGCCCACTCTGACCGCGTTCCTGAGCTCTCTCTGAGTCTCT 8890  
Db CACCACTTGAACCAAGCCCACTCTGACCGCGTTCCTGAGCTCTCTCTGAGTCTCT 8890  
QY 1679 CCGCGCGCAGCTGT 1738  
Db CCGCGCGCAGCTGT 1738  
QY 8891 CCGCGCGCAGCTGT 8950  
Db CCGCGCGCAGCTGT 8950  
QY 1739 AAATGGAATCATTTGACTGTGAAAAGCTGAGTACTCCTCTCTGAGCGAGCCCTCAACC 1798  
Db AAATGGAATCATTTGACTGTGAAAAGCTGAGTACTCCTCTCTGAGCGAGCCCTCAACC 1798  
QY 8951 AAATGGAATCATTTGACTGTGAAAAGCTGAGTACTCCTCTCTGAGCGAGCCCTCAACC 9010  
Db AAATGGAATCATTTGACTGTGAAAAGCTGAGTACTCCTCTCTGAGCGAGCCCTCAACC 9010  
QY 1799 AGTTCAGATCCAGCGCTCCACCGGAGCGCTTCTCTCTGCTGCCAACAAGGTTTC 1858  
Db AGTTCAGATCCAGCGCTCCACCGGAGCGCTTCTCTCTGCTGCCAACAAGGTTTC 1858  
QY 9011 AGTTCAGATCCAGCGCTCCACCGGAGCGCTTCTCTCTGCTGCCAACAAGGTTTC 9070  
Db AGTTCAGATCCAGCGCTCCACCGGAGCGCTTCTCTCTGCTGCCAACAAGGTTTC 9070  
QY 1859 CGTGGCTGTTTGTGAGCTTAACATTTGACCTTCCGCAATTTGAGCTTCAAGCAA 1918  
Db CGTGGCTGTTTGTGAGCTTAACATTTGACCTTCCGCAATTTGAGCTTCAAGCAA 1918  
QY 9071 CGTGGCTGTTTGTGAGCTTAACATTTGACCTTCCGCAATTTGAGCTTCAAGCAA 9130  
Db CGTGGCTGTTTGTGAGCTTAACATTTGACCTTCCGCAATTTGAGCTTCAAGCAA 9130  
QY 1919 TTGCAACAGCGTGGGAGGAGCAGAGCAGAGCTGCTTTTATAGTCTCCATTTCAAG 1978  
Db TTGCAACAGCGTGGGAGGAGCAGAGCAGAGCTGCTTTTATAGTCTCCATTTCAAG 1978  
QY 9131 TTGCAACAGCGTGGGAGGAGCAGAGCAGAGCTGCTTTTATAGTCTCCATTTCAAG 9190  
Db TTGCAACAGCGTGGGAGGAGCAGAGCAGAGCTGCTTTTATAGTCTCCATTTCAAG 9190  
QY 1979 AGATTAACACCGAGACTTGGAGGAGCAGATGAGCAAGGCGCCACTTTGTCCCTTA 2038  
Db AGATTAACACCGAGACTTGGAGGAGCAGATGAGCAAGGCGCCACTTTGTCCCTTA 2038  
QY 9191 AGATTAACACCGAGACTTGGAGGAGCAGATGAGCAAGGCGCCACTTTGTCCCTTA 9250  
Db AGATTAACACCGAGACTTGGAGGAGCAGATGAGCAAGGCGCCACTTTGTCCCTTA 9250  
QY 2039 GCMAATTCAAGGATCACTTCACTTAACCAAGGCTGCTTCACTGTGCTGCTTCTCA 2098  
Db GCMAATTCAAGGATCACTTCACTTAACCAAGGCTGCTTCACTGTGCTGCTTCTCA 2098  
QY 9251 GCMAATTCAAGGATCACTTCACTTAACCAAGGCTGCTTCACTGTGCTGCTTCTCA 9310  
Db GCMAATTCAAGGATCACTTCACTTAACCAAGGCTGCTTCACTGTGCTGCTTCTCA 9310  
QY 2099 AGCATTTATTAAGACCTTAAGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2158  
Db AGCATTTATTAAGACCTTAAGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2158  
QY 9311 AGCATTTATTAAGACCTTAAGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 9370  
Db AGCATTTATTAAGACCTTAAGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 9370  
QY 2159 CCCAGTCTGAGCTGCG 2218  
Db CCCAGTCTGAGCTGCG 2218  
QY 9371 CCCAGTCTGAGCTGCG 9430  
Db CCCAGTCTGAGCTGCG 9430  
QY 2219 GGGAGGGGGGAGAGAGAGTGGAGCCAGAACTGAGCTGGAGGGATCCGACAGAAA 2278  
Db GGGAGGGGGGAGAGAGAGTGGAGCCAGAACTGAGCTGGAGGGATCCGACAGAAA 2278  
QY 9431 GGGAGGGGGGAGAGAGAGTGGAGCCAGAACTGAGCTGGAGGGATCCGACAGAAA 9490  
Db GGGAGGGGGGAGAGAGAGTGGAGCCAGAACTGAGCTGGAGGGATCCGACAGAAA 9490  
QY 2279 GCTCAGGGGGGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2338  
Db GCTCAGGGGGGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2338  
QY 9491 GCTCAGGGGGGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9550  
Db GCTCAGGGGGGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9550  
QY 2339 TCAGGCAATGGGTGTTTGAACCAATTAACCAAGGCTTGTATCAGCTCTTAACAG 2398  
Db TCAGGCAATGGGTGTTTGAACCAATTAACCAAGGCTTGTATCAGCTCTTAACAG 2398  
QY 9551 TCAGGCAATGGGTGTTTGAACCAATTAACCAAGGCTTGTATCAGCTCTTAACAG 9610  
Db TCAGGCAATGGGTGTTTGAACCAATTAACCAAGGCTTGTATCAGCTCTTAACAG 9610  
QY 2399 TATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2458  
Db TATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2458  
QY 9611 TATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 9670  
Db TATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 9670  
QY 2459 GTGATCTTCTAATTTGGCTTCTTGTGAGGTTGAGAGGGTAAAGTAACTTCTCAAGTCAAC 2518  
Db GTGATCTTCTAATTTGGCTTCTTGTGAGGTTGAGAGGGTAAAGTAACTTCTCAAGTCAAC 2518  
QY 9671 GTGATCTTCTAATTTGGCTTCTTGTGAGGTTGAGAGGGTAAAGTAACTTCTCAAGTCAAC 9730  
Db GTGATCTTCTAATTTGGCTTCTTGTGAGGTTGAGAGGGTAAAGTAACTTCTCAAGTCAAC 9730  
QY 2519 ACAGCAAGTCTGTGGGTGGAG 2578  
Db ACAGCAAGTCTGTGGGTGGAG 2578

```

Db      9731 ACAGCAAGCTGTGGTGGGAGAGCAAGCTAGGCGTGGGCAATTCAGTACATACACAGAT 9790
Oy      2579 GTCCTCCCTCTTGTGATGCTTGGCCCTTGAAGGCTTTGGGAATCTTGCTT 2638
Db      9791 GTCCTCCCTCTTGTGATGCTTGGCCCTTGAAGGCTTTGGGAATCTTGCTT 9850
Oy      2639 CAACCCCTCCCTAGATCAGTGTGAGGGTCCCTGTAGATTTGTGTACACCATGCCCCA 2698
Db      9851 CAACCCCTCCCTAGATCAGTGTGAGGGTCCCTGTAGATTTGTGTACACCATGCCCCA 9910
Oy      2699 TGTATATACAGTACACACAGATGTACACACAGATGTACATGCTCCAGCCCACTCT 2758
Db      9911 TGTATATACAGTACACACAGATGTACACACAGATGTACATGCTCCAGCCCACTCT 9970
Oy      2759 GCATACCTGACCTGACACCCCAAGCTTGGCCCTGCTGCTGTGCTCTCAAGCAGCAG 2818
Db      9971 GCATACCTGACCTGACACCCCAAGCTTGGCCCTGCTGCTGTGCTCTCAAGCAGCAG 10030
Oy      2819 CTCACACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2878
Db      10031 CTCACACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10090
Oy      2879 TACCTTGCTGACACCGGTGTGTGGCCGCTCTCACCCAGGACACGCCCCGACCATGGA 2938
Db      10091 TACCTTGCTGACACCGGTGTGTGGCCGCTCTCACCCAGGACACGCCCCGACCATGGA 10150
Oy      2939 TCTCCGTGACACTATCATTAATAAGTGGGTTTGTACAA 2978
Db      10151 TCTCCGTGACACTATCATTAATAAGTGGGTTTGTACAA 10190

```

```

RESULT 8
HSS69D19 114771 bp DNA linear PRI 05-JUN-2003
LOCUS Human DNA sequence from clone RP4-569D19 on chromosome 22q13.1,
DEFINITION complete sequence.
ACCESSION AL022334
VERSION AL022334.1 GI:3947839
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 114771)
AUTHORS Williams, S.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Dec 2, 1998 this sequence version replaced gi:3945463.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
http://www.sanger.ac.uk/Projects/C_elegans/wormpep -----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk

```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>  
RP4-569D19 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pCYPAC2  
IMPORTANT: This sequence is not the entire insert of clone RP4-569D19. It may be shorter because we sequence overlapping sections only once, except for a short overlap.  
The true left end of clone RP4-569D19 is at 1 in this sequence. The true right end of clone RP5-824119 is at 31286 in this sequence.

## FEATURES

```

source
repeat_region
  1..114771
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="22"
  /map="q13.1"
  /clone="RP4-569D19"
  /clone_11b="RPCI-4"
  142..175
  /note="MIR3 repeat: matches 162..195 of consensus"
  277..290
  /note="2.3 copies 6 mer GAGAA 28% conserved"
  418..427
  /note="3.3 copies 3 mer CTC 20% conserved"
  514..524
  /note="2.2 copies 5 mer CAGCA 22% conserved"
  756..765
  /note="2.5 copies 4 mer CCCA 20% conserved"
  /complement(805..1011)
  /note="MIR repeat: matches 50..258 of consensus"
  1012..1253
  /note="AluSc repeat: matches 61..304 of consensus"
  /complement(1466..2088)
  /note="L2 repeat: matches 2487..3183 of consensus"
  2302..2611
  /note="AluSg repeat: matches 1..309 of consensus"
  2616..2630
  /note="3.8 copies 4 mer AAC 30% conserved"
  2631..2706
  /note="2.7 copies 6 mer GTGCA 25% conserved"
  2748..3041
  /note="L2 repeat: matches 3033..3313 of consensus"
  3546..3871
  /note="AluSx repeat: matches 1..312 of consensus"
  3886..4117
  /note="MIR repeat: matches 13..237 of consensus"
  /complement(4206..4486)
  /note="AluYb repeat: matches 1..282 of consensus"
  /complement(4747..4858)
  /note="MIR repeat: matches 47..163 of consensus"
  4882..4959
  /note="L2 repeat: matches 3187..3266 of consensus"
  5079..5101
  /note="1.9 copies 12 mer AATCATCATG 46% conserved"
  5080..5103
  /note="6.0 copies 4 mer ATTC 30% conserved"
  5108..5368
  /note="match: GSS: Em:A2520278"
  5219..5229
  /note="2.8 copies 4 mer TCTT 22% conserved"
  5324..5334
  /note="12.2 copies 5 mer CTCTC 22% conserved"
  5493..5545
  /note="2.1 copies 25 mer GAGGGAGAGCTTTATTCACGCA 88% conserved"
  5693..5990
  /note="AluY repeat: matches 1..300 of consensus"

```

repeat\_region 5999. .6140  
/note="AluSg/x repeat: matches 1. .136 of consensus"  
repeat\_region 6147. .6202  
/note="MIR repeat: matches 151. .205 of consensus"  
repeat\_region complement(6238. .6292)  
/note="MIR3 repeat: matches 147. .203 of consensus"  
repeat\_region 6330. .6339  
/note="10.0 copies 1 mer T 20% conserved"  
repeat\_region 6342. .6352  
/note="2.2 copies 5 mer GTGGG 22% conserved"  
repeat\_region 6453. .7182  
/note="L1M45 repeat: matches 4129. .4859 of consensus"  
repeat\_region complement(7294. .7605)  
/note="AluX repeat: matches 1. .312 of consensus"  
repeat\_region 7630. .7640  
/note="2.8 copies 4 mer ATTA 22% conserved"  
repeat\_region 7786. .8087  
/note="AluSg repeat: matches 6. .307 of consensus"  
repeat\_region 8748. .8909  
/note="AluYa5a2 repeat: matches 149. .310 of consensus"  
repeat\_region 8910. .8965  
/note="28.0 copies 2 mer AT 112% conserved"  
repeat\_region 9217. .9231  
/note="3.0 copies 5 mer GCTT 21% conserved"  
misc\_feature complement(9458. .9902)  
/note="match: GSS: Em:A0222044"  
misc\_feature complement(9546. .9904)  
/note="match: GSS: Em:B57571"  
repeat\_region 10091. .10506  
/note="L1M45 repeat: matches 4845. .5273 of consensus"  
repeat\_region 10508. .10592  
/note="AluY/FLAM repeat: matches 1. .88 of consensus"  
repeat\_region 10593. .11608  
/note="L1M45 repeat: matches 5265. .6298 of consensus"  
repeat\_region 11640. .11653  
/note="14.0 copies 1 mer A 28% conserved"  
repeat\_region complement(11931. .12080)  
/note="L1M1 repeat: matches 6010. .6161 of consensus"  
repeat\_region 12092. .12108  
/note="2.4 copies 7 mer TGTTTC 25% conserved"  
repeat\_region 12150. .12164  
/note="7.5 copies 2 mer AG 23% conserved"  
repeat\_region complement(12342. .12784)  
/note="L1M1 repeat: matches 5573. .6011 of consensus"  
repeat\_region complement(13431. .13736)  
/note="AluX repeat: matches 1. .306 of consensus"  
repeat\_region 13794. .14127  
/note="LTR9 repeat: matches 1. .270 of consensus"  
repeat\_region 14128. .14145  
/note="AluX repeat: matches 1. .289 of consensus"  
repeat\_region 14416. .14737  
/note="LTR9 repeat: matches 270. .612 of consensus"  
repeat\_region 14741. .14756  
/note="2.3 copies 7 mer ACCTGCT 23% conserved"  
repeat\_region complement(14909. .15033)  
/note="FLAM C repeat: matches 2. .126 of consensus"  
repeat\_region 15054. .15074  
/note="10.5 copies 2 mer AC 33% conserved"  
repeat\_region 15083. .15640  
/note="L2 repeat: matches 2715. .3306 of consensus"  
repeat\_region complement(15622. .15787)  
/note="MIR repeat: matches 85. .256 of consensus"  
repeat\_region 15851. .15860  
/note="2.5 copies 4 mer AATA 20% conserved"  
repeat\_region 16065. .16079  
/note="2.5 copies 6 mer AGCCC 30% conserved"  
repeat\_region 16383. .16392  
/note="2.5 copies 4 mer TTA 20% conserved"  
repeat\_region 16487. .16510  
/note="12.0 copies 2 mer CT 48% conserved"  
repeat\_region 16510. .16526  
/note="8.5 copies 2 mer TG 34% conserved"  
repeat\_region 16553. .16686

misc\_feature 16658. .16890  
/note="match: GSS: Em:B13882"  
repeat\_region 16850. .16976  
/note="MIR repeat: matches 16. .148 of consensus"  
repeat\_region 17399. .17409  
/note="2.2 copies 5 mer TAATT 22% conserved"  
repeat\_region 17416. .17430  
/note="15.0 copies 1 mer A 21% conserved"  
repeat\_region 17734. .17806  
/note="LTR16A1 repeat: matches 12. .84 of consensus"  
repeat\_region 17842. .18145  
/note="LTR16A1 repeat: matches 147. .456 of consensus"  
repeat\_region 18411. .18441

Query Match 79.5%; Score 2449; DB 9; Length 114771;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 479 CAGGGATGTTTATCTGTTGTTCAAGCTTGATTAACGGGATCTTGATGATGATCA 538  
|||  
Db 57923 CAGGGATGTTTATCTGTTGTTCAAGCTTGATTAACGGGATCTTGATGATGATCA 57982

QY 539 AGCCCTTCAGAAACAGATCTGAGGTCAAGTCTGCTGAAGAACAAAGCAAGAGAG 598  
|||  
Db 57983 AGCCCTTCAGAAACAGATCTGAGGTCAAGTCTGCTGAAGAACAAAGCAAGAGAG 58042

QY 599 CGGCGAGCTGCCCATGCTCATCTGTGGCAAGAAAGCAAGCAAGCGAGCTGTGCCGC 658  
|||  
Db 58043 CGGCGAGCTGCCCATGCTCATCTGTGGCAAGAAAGCAAGCAAGCGAGCTGTGCCGC 58102

QY 659 AGGTGCCACACAGAGCGAGGCTGCTGTGTGGGCGAGAGACTCCGCTCACTTCG 718  
|||  
Db 58103 AGGTGCCACACAGAGCGAGGCTGCTGTGTGGGCGAGAGACTCCGCTCACTTCG 58162

QY 719 AGGTGTGGCCAGAAAGAACCAACAGTGAACAGATGTTCTAGTCTCTTCAGCATGG 778  
|||  
Db 58163 AGGTGTGGCCAGAAAGAACCAACAGTGAACAGATGTTCTAGTCTCTTCAGCATGG 58222

QY 779 CCAAGCTGCCACAGATGAGCCGCCCTGCTATGCCAGATCTCGTGAGTACGATG 838  
|||  
Db 58223 CCAAGCTGCCACAGATGAGCCGCCCTGCTATGCCAGATCTCGTGAGTACGATG 58282

QY 839 AGCGCTTCACACCCAGAGCCCTTGTCAGTGGCCGCTCAAGAGATGAGAGCCATGAGCA 898  
|||  
Db 58283 AGCGCTTCACACCCAGAGCCCTTGTCAGTGGCCGCTCAAGAGATGAGAGCCATGAGCA 58342

QY 899 TGGTCTGCGCTTGCCGCCGCCGCCAGCGTCAACAGTGAACCTCAATCAATCAAGGCCA 958  
|||  
Db 58343 TGGTCTGCGCTTGCCGCCGCCGCCAGCGTCAACAGTGAACCTCAATCAATCAAGGCCA 58402

QY 959 AGGTCTTTGGGAAAGGCGAGCCCGCTGAGAGGAGCAAGTGCATCATGAGCGAGG 1018  
|||  
Db 58403 AGGTCTTTGGGAAAGGCGAGCCCGCTGAGAGGAGCAAGTGCATCATGAGCGAGG 58462

QY 1019 ATGTGGGGGGGGCTTGCGCAAGTCTTCAAGGAGGTGGCCCAATGCCACTGTGGC 1078  
|||  
Db 58463 ATGTGGGGGGGGCTTGCGCAAGTCTTCAAGGAGGTGGCCCAATGCCACTGTGGC 58522

QY 1079 CATCTCCCAACGAGGCCCGGAGAGTCTTGTCAAGACCTTAGAGCAAGACTGGA 1138  
|||  
Db 58523 CATCTCCCAACGAGGCCCGGAGAGTCTTGTTCACAGACCTTAGAGCAAGACTGGA 58582

QY 1139 GGGCCCCGGGGGCTTGCGCTCGCAAGTCTTCAAGGAGGTGGCCCAATGCCACTGTGGC 1198  
|||  
Db 58583 GGGCCCCGGGGGCTTGCGCTCGCAAGTCTTCAAGGAGGTGGCCCAATGCCACTGTGGC 58642

QY 1199 TTGTCCACAGCTCTTGATGATTTCAATCTCTGTGGAGAGCACATCTCTCAGGCTC 1258  
|||  
Db 58643 TTGTCCACAGCTCTTGATGATTTCAATCTCTGTGGAGAGCACATCTCTCAGGCTC 58702

QY 1259 AAGAGTTAGGCAAGACTCAAGTTAACCCTTCTCTGCGGTTGAGAGAAATGTTGAT 1318  
|||









2278 AGCTCAGGCGGGCTCTCTCCCTGTCGCCGGATTGGCTATGCTGGGTACCAACATGTA 2337  
|  
|  
|  
Db 1982 AGCTCAGGCGGGCTCTCTCCCTGTCGCCGGATTGGCTATGCTGGGTACCAACATGTA 2041  
|  
|  
|  
Qy 2338 CTCAGGCGGGCTCTCTCCCTGTCGCCGGATTGGCTATGCTGGGTACCAACATGTA 2397  
|  
|  
|  
Db 2042 CTCAGGCGGGCTCTCTCCCTGTCGCCGGATTGGCTATGCTGGGTACCAACATGTA 2101  
|  
|  
|  
Qy 2398 GTATATTTTGTATTTTAAATCTCTTAAACATTTGAAGTTTAAAGGCTTAAAGAACCTT 2457  
|  
|  
|  
Db 2102 GTATATTTTGTATTTTAAATCTCTTAAACATTTGAAGTTTAAAGGCTTAAAGAACCTT 2161  
|  
|  
|  
Qy 2458 AGTATCTTCTATTTGGCTCTTCTGAGTTTCAAGAGGTTAATTAATCTCTCAAGTTCA 2517  
|  
|  
|  
Db 2162 AGTATCTTCTATTTGGCTCTTCTGAGTTTCAAGAGGTTAATTAATCTCTCAAGTTCA 2221  
|  
|  
|  
Qy 2518 CACAGCAAGTCTGTGGTGGCAGAGCAAGCTAGCGTGGCATTCAGTACATACCAAGA 2577  
|  
|  
|  
Db 2222 CACAGCAAGTCTGTGGTGGCAGAGCAAGCTAGCGTGGCATTCAGTACATACCAAGA 2281  
|  
|  
|  
Qy 2578 TGTGCTCCCTCTCTTAAATGCTTGGCCCTTGGGGCTTCAAGGCTTTGGAGATCTTGTCC 2637  
|  
|  
|  
Db 2282 TGTGCTCCCTCTCTTAAATGCTTGGCCCTTGGGGCTTCAAGGCTTTGGAGATCTTGTCC 2341  
|  
|  
|  
Qy 2638 TCAACCTCTCCCTAGATGATCTGTGAGGGTCCCTGATATTTGTATACCATGCTCC 2697  
|  
|  
|  
Db 2342 TCAACCTCTCCCTAGATGATCTGTGAGGGTCCCTGATATTTGTATACCATGCTCC 2401  
|  
|  
|  
Qy 2698 ATGATATTAACAAGTACACAGATGTACACAGATGTACACATGTCTCAGCCCACTC 2757  
|  
|  
|  
Db 2402 ATGATATTAACAAGTACACAGATGTACACAGATGTACACATGTCTCAGCCCACTC 2461  
|  
|  
|  
Qy 2758 TGCATACCTGCACTGCAACCCAGCTTGGCCCTTGGGGCTTGTGCTTCAAGAGCA 2817  
|  
|  
|  
Db 2462 TGCATACCTGCACTGCAACCCAGCTTGGCCCTTGGGGCTTGTGCTTCAAGAGCA 2521  
|  
|  
|  
Qy 2818 GCTTCAACCTGCTGTGCTTCCCTTCCCAACCACTGCTGAGGCTTGTGAGAGCAAG 2877  
|  
|  
|  
Db 2522 GCTTCAACCTGCTGTGCTTCCCTTCCCAACCACTGCTGAGGCTTGTGAGAGCAAG 2581  
|  
|  
|  
Qy 2878 GTACTTGGCTGCAACCGGTGTGTGGCCGCTCTGACCCAGAGCAGCCGCCACATGG 2937  
|  
|  
|  
Db 2582 GTACTTGGCTGCAACCGGTGTGTGGCCGCTCTGACCCAGAGCAGCCGCCACATGG 2641  
|  
|  
|  
Qy 2938 ATCTCCGTGTACTATCAATTAAGTGGGTTGTGTACAAAAAATTTTTTTTTT 2995  
|  
|  
|  
Db 2642 ATCTCCGTGTACTATCAATTAAGTGGGTTGTGTACAAAAAATTTTTTTTTT 2699  
|  
|  
|  
RESULT 10  
AX597851 1022 bp DNA linear PAT 14-FEB-2003  
LOCUS AX597851  
DEFINITION Sequence 125 from Patent WO0244994.  
ACCESSION AX597851  
VERSION AX597851.1 GI:28398024  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Brower, A., Brow, M.A., Gracauer, R.F., Fors, L., Granske, R., de arruda  
Indig, M., Kurensky, D., Luedtke, C., Lukowiak, A.A., Lyamichiev, V.,  
Neri, B.P., Reimer, N.D., Roeven, R.T., Skrzybczynski, Z., Ziatno, W.A.,  
Comerford, J., Stump, S. and Vilegus, D.D.  
TITLE System and method for detection assay production and sale  
JOURNAL Patent: WO 0244994-A 125 06-JUN-2002;  
THIRD WAVE TECHNOLOGIES, INC. (US)  
FEATURES  
SOURCE Location/Qualifiers  
1. 1022  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 27.8%; Score 858; DB 6; Length 1022;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 978; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 708 CGCTTACTTGAAGTGTGCGCCAAAGAAACACCAACGTTGAGAGATGTTCTACGTC 767  
|  
|  
|  
Db 44 CGCTTACTTGAAGTGTGCGCCAAAGAAACACCAACGTTGAGAGATGTTCTACGTC 103  
|  
|  
|  
Qy 768 CTTCAGATGCGCAAGCTGCAACAGATGAGCCCGCTTCGATGCAAGATCTTCCT 827  
|  
|  
|  
Db 104 CTTCAGATGCGCAAGCTGCAACAGATGAGCCCGCTTCGATGCAAGATCTTCCT 163  
|  
|  
|  
Qy 828 GCATGACGTTGAGAGCTTCCACCCAGGCTTCTGATGAGCGCGCTCAAGAGATGGA 887  
|  
|  
|  
Db 164 GCATGAGGTTGAGAGCTTCCACCCAGGCTTCTGATGAGCGCGCTCAAGAGATGGA 223  
|  
|  
|  
Qy 888 CGCTTATGAGATGATCTGCGCTTCCGCGCGCCCGCCAGCGTCAACAGTACCTCAAGTA 947  
|  
|  
|  
Db 224 CGCTTATGAGATGATCTGCGCTTCCGCGCGCCCGCCAGCGTCAACAGTACCTCAAGTA 283  
|  
|  
|  
Qy 948 CATCAAGCGCAAGTCTTGGGAAAGCGCCGCTGAGAGGAGCAAGTGCACATCCA 1007  
|  
|  
|  
Db 284 CATCAAGCGCAAGTCTTGGGAAAGCGCCGCTGAGAGGAGCAAGTGCACATCCA 343  
|  
|  
|  
Qy 1008 GTGAGCGAGGATGCTGAGGCGGGGCTTGGCCAGTGCCTTCAGAGAGTGGCCAGATG 1067  
|  
|  
|  
Db 344 GTGAGCGAGGATGCTGAGGCGGGGCTTGGCCAGTGCCTTCAGAGAGTGGCCAGATG 403  
|  
|  
|  
Qy 1068 CCCAGTGTGCGCATCTCCCAACGAGGCGCCGAGAGAGCTTGTTCACAGACTTAAAGG 1127  
|  
|  
|  
Db 404 CCCAGTGTGCGCATCTCCCAACGAGGCGCCGAGAGAGCTTGTTCACAGACTTAAAGG 463  
|  
|  
|  
Qy 1128 ACCAGATGAGAGCGCCCGGAGCGCTGCGCATTTGCTGCTTCAAGCTTT 1187  
|  
|  
|  
Db 464 ACCAGATGAGAGCGCCCGGAGCGCTGCGCATTTGCTGCTTCAAGCTTT 523  
|  
|  
|  
Qy 1188 CTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCA-TCTCTCTGTGGAGAGCAAT 1246  
|  
|  
|  
Db 524 CTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCA-TCTCTCTGTGGAGAGCAAT 583  
|  
|  
|  
Qy 1247 CTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCA-TCTCTCTGTGGAGAGCAAT 1306  
|  
|  
|  
Db 584 CTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTCA-TCTCTCTGTGGAGAGCAAT 643  
|  
|  
|  
Qy 1307 AGAATGTTGATGCAAGAGGAGGATGAGATGCTGCGCTCATATGAGAGCTCTGAGCAAG 1366  
|  
|  
|  
Db 644 AGAATGTTGATGCAAGAGGAGGATGAGATGCTGCGCTCATATGAGAGCTCTGAGCAAG 703  
|  
|  
|  
Qy 1367 CTTCAGAGTAAAGAGCAGAGAGGCGCAGATGAGAAAGTCTTCTCTCTGAGCATTA 1426  
|  
|  
|  
Db 704 CTTCAGAGTAAAGAGCAGAGAGGCGCAGATGAGAAAGTCTTCTCTCTGAGCATTA 763  
|  
|  
|  
Qy 1427 CACCAGCTTGTGTTGGGTGGAGCTGGAGAACTTCTCCAGAGCTGCAACTTTAC 1486  
|  
|  
|  
Db 764 CACCAGCTTGTGTTGGGTGGAGCTGGAGAACTTCTCTCCAGAGCTGCAACTTTAC 823  
|  
|  
|  
Qy 1547 CAGCTGCGCTTGTACATTTGAGCCAGTGAATCTGTGTGTAAGGGGGCTGGCCACACTTC 1606  
|  
|  
|  
Db 884 CAGCTGCGCTTGTACATTTGAGCCAGTGAATCTGTGTGTAAGGGGGCTGGCCACACTTC 943  
|  
|  
|  
Qy 1607 CTGAGCAGGCGCAACCACTTGAACAGGCGCACTCTGACCGGTTCTCAGGCTCTC 1666  
|  
|  
|  
Db 944 CTGAGCAGGCGCAACCACTTGAACAGGCGCACTCTGACCGGTTCTCAGGCTCTC 1003  
|  
|  
|  
Qy 1667 TCCTAGGTCCCTCGCCCG 1685  
|  
|  
|  
Db 1004 TCCTAGGTCCCTCGCCCG 1022

RESULT 11  
LOCUS AR412679 478 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 316 from patent US 6639063.  
ACCESSION AR412679  
VERSION AR412679.1 GI:40167789  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 478)  
AUTHORS Edwards, J. -B.D.M., Jobert, S. and Giordano, J. -Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 316 28-OCT-2003;  
FEATURES  
source  
location/Qualifiers  
1..478  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 8.2%; Score 253; DB 6; Length 478;  
Best Local Similarity 99.7%; Pred. No. 2e-136;  
Matches 373; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 2424 AACATATTGAAGTTTAAAGGCGCTTAAGAACTTATGATCTTCTATTGGCTCTTCTGA 2483  
DB 1 AACATATTGAAGTTTAAAGGCGCTTAAGAACTTATGATCTTCTATTGGCTCTTCTGA 60  
QY 2484 GGTTCAGAGAGGGTAAGTAATCTCTCCAGGTCAACAGCAAGCTGTGGGTGAGAGAAG 2543  
DB 61 GGTTCAGAGAGGGTAAGTAATCTCTCCAGGTCAACAGCAAGCTGTGGGTGAGAGAAG 120  
QY 2544 CAAGCTACGCTGGGCAATTCAGTACATACCAAGATGCTCTCTCTTGTATGCTTGGCC 2603  
DB 121 CAAGCTACGCTGGGCAATTCAGTACATACCAAGATGCTCTCTCTTGTATGCTTGGCC 180  
QY 2604 CTTGGGGGCTTCAGGGCTTTGGAGATCTGTCTCTCAACCTCTCTCCCTAGATCACTCTGT 2663  
DB 181 CTTGGGGGCTTCAGGGCTTTGGAGATCTGTCTCTCAACCTCTCTCCCTAGATCACTCTGT 239  
QY 2664 GAGGCTCCCTGTATATTTGTGTATACACCATGCCATGTATATACAAAGTACACAGATGT 2723  
DB 240 GAGGCTCCCTGTATATTTGTGTATACACCATGCCATGTATATACAAAGTACACAGATGT 299  
QY 2724 ACAACAGATGTATACATGCTCCAGCCCAAGCTGTGATACCTGACCTGACCCAGCC 2783  
DB 300 ACAACAGATGTATACATGCTCCAGCCCAAGCTGTGATACCTGACCTGACCCAGCC 359  
QY 2784 TTGGCCCTGCTG 2797  
DB 360 TTGGCCCTGCTG 373

RESULT 12  
LOCUS BD108232 478 bp DNA linear PAT 18-SEP-2002  
DEFINITION EST and encoded human protein.  
ACCESSION BD108232  
VERSION BD108232.1 GI:23203050  
KEYWORDS JP 2002010789-A/309.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 478)  
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 309 15-JAN-2002;  
GENSET CORP  
OS Homo sapiens (human)  
PN JP 2002010789-A/309

PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC Von Heijne matrix  
CC score 4  
CC seq ALAPAXVCAQSSS/SN  
CC n=a, y, c or t  
FH Key location/Qualifiers  
FT sig\_peptide 171..476  
FT sig\_peptide 171..395  
FT misc feature 463.  
location/Qualifiers  
1..478  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 8.2%; Score 253; DB 6; Length 478;  
Best Local Similarity 99.7%; Pred. No. 2e-136;  
Matches 373; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 2424 AACATATTGAAGTTTAAAGGCGCTTAAGAACTTATGATCTTCTATTGGCTCTTCTGA 2483  
DB 1 AACATATTGAAGTTTAAAGGCGCTTAAGAACTTATGATCTTCTATTGGCTCTTCTGA 60  
QY 2484 GGTTCAGAGAGGGTAAGTAATCTCTCCAGGTCAACAGCAAGCTGTGGGTGAGAGAAG 2543  
DB 61 GGTTCAGAGAGGGTAAGTAATCTCTCCAGGTCAACAGCAAGCTGTGGGTGAGAGAAG 120  
QY 2544 CAAGCTACGCTGGGCAATTCAGTACATACCAAGATGCTCTCTCTTGTATGCTTGGCC 2603  
DB 121 CAAGCTACGCTGGGCAATTCAGTACATACCAAGATGCTCTCTCTTGTATGCTTGGCC 180  
QY 2604 CTTGGGGGCTTCAGGGCTTTGGAGATCTGTCTCTCAACCTCTCTCCCTAGATCACTCTGT 2663  
DB 181 CTTGGGGGCTTCAGGGCTTTGGAGATCTGTCTCTCAACCTCTCTCCCTAGATCACTCTGT 239  
QY 2664 GAGGCTCCCTGTATATTTGTGTATACACCATGCCATGTATATACAAAGTACACAGATGT 2723  
DB 240 GAGGCTCCCTGTATATTTGTGTATACACCATGCCATGTATATACAAAGTACACAGATGT 299  
QY 2724 ACAACAGATGTATACATGCTCCAGCCCAAGCTGTGATACCTGACCTGACCCAGCC 2783  
DB 300 ACAACAGATGTATACATGCTCCAGCCCAAGCTGTGATACCTGACCTGACCCAGCC 359  
QY 2784 TTGGCCCTGCTG 2797  
DB 360 TTGGCCCTGCTG 373

RESULT 13  
LOCUS HSM801346 2355 bp mRNA linear PRI 18-FEB-2000  
DEFINITION Homo sapiens mRNA; CDNA DKFZp434M0317 (from clone DKFZp434M0317).  
ACCESSION AL133075  
VERSION AL133075.1 GI:6453518  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2355)  
AUTHORS Bloecher, H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and  
Wiemann, S.  
TITLE Direct Submission  
JOURNAL Submitted (15-NOV-1999) MIPS, Am Kioferpititz 18a, D-82152

COMMENT Martinried, GERMANY  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de, Braunschweig/Germany) within the CDNA sequencing consortium of the German Genome Project.  
 This clone (DKFZ434M0317) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cdna/>.

FEATURES  
 source  
 1..2355  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZ434M0317"  
 /tissue\_type="testis"  
 /clone\_lib="434 (synonym: htes3). Vector pSPORT1; host DH10B; sites NotI + SalI"  
 /dev\_stage="adult"  
 2217..2222

ORIGIN  
 polyA\_signal  
 polyA\_site  
 2217

Query Match 3.6%; Score 111; DB 9; Length 2355;  
 Best Local Similarity 100.0%; Freq. No. 7.2e-53;  
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2972 TTACCAA 3031  
 |||  
 Db 2236 TTACCAA 2295  
 |||

Qy 3032 AA 3082  
 |||  
 Db 2296 AA 2346  
 |||

RESULT 14  
 BC027800  
 LOCUS  
 DEFINITION Mus musculus PTK7 protein tyrosine kinase 7, mRNA (cDNA clone IMAGE:5151185), partial cds.  
 ACCESSION BC027800  
 VERSION BC027800.1 GI:20380016  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2470)  
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F., Diatchenko,L., Marisina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uebli,T.B., Toshlyuk,S., Carninci,P., Prange,C., Raha,S.S., Loguettano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Vialation,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smailus,D.E., Scherch,A., Schein,J.B., Jones,S.J., and Marra,M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257

PUBMED 12477932  
 2 (bases 1 to 2470)  
 Strausberg,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-@mail.nih.gov](mailto:cgabs-@mail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Akhter N., Ayala K., Beckstrom-Sternberg S.M., Benjamin B., Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S., Dietrich N.L., Granite S., Guan X., Gupta J., Haghighi P., Hansen N., Ho S.-L., Karlins E., Kwong P., Latic P., Legaspi R., Maduro Q.L., Mastello C., Maskeri B., Mastrian S.D., McCloskey J.C., McDowell J., Pearson R., Stantropop S., Thomas P.J., Touchman J.W., Turgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L., Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>  
 Series: IRAC Plate: 56 Row: a Column: 12  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES  
 source  
 1..2470  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5151185"  
 /tissue\_type="Mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue."  
 /clone\_lib="NCI CGAP Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 <1..2470  
 /gene="Ptk7"  
 /db\_xref="LocusID:71461"  
 /db\_xref="MGI:1918711"  
 <1..1426  
 /gene="Ptk7"  
 /codon\_start=2  
 /product="Ptk7 protein"  
 /protein\_id="AAH27800.1"  
 /db\_xref="GI:20380017"  
 /db\_xref="LocusID:71461"  
 /db\_xref="MGI:1918711"  
 /translation="HSGTIVYQHTALRCEAQGPXPLIOWKGRDIIDPTLGR MHIFQNSLIVHIVAPEDSSGYSIGNSINHTPEAPLIVVQPVWDESGSPPP KMIQITGLGVGAIVVYIIVLIGMFCKRCKAKLOKOPGEPEMECNGSLPN GQPSARIGEVALLTSLGSGPATPKKSHAGDMHFPASLOPITTTGSRGVEFLAK AGVEBAGETTLVAVSLGSRDECOQIDPREYVEMREKLNHNAVYRLIGCREAEPHY MVAEYDGLDKLPFLRSKKRDKLSQPISTYOKVALCSQVALGMSLNNRNVKHO LVARNCILISQVQKVSALGSDVNSFYHROAVVPLRMSPEAVLSDGPESTKSD VAFGVLMVFPFHGEMPHGQADDEVTLADLQGKARLPOPEGCPSTLYRLMORCWAP NPDORPFSFIATSLGSPADSKQ"

misc\_feature  
 17..256  
 /gene="Ptk7"  
 /note="IG-like; Region: Immunoglobulin like"  
 /db\_xref="CDD:cd00097"  
 581..1402  
 /gene="Ptk7"

misc\_feature  
 17..256  
 /gene="Ptk7"  
 /note="IG-like; Region: Immunoglobulin like"  
 /db\_xref="CDD:cd00097"  
 581..1402  
 /gene="Ptk7"



Tue Jun 22 10:07:19 2004

us-09-778-963a-1.011.rge

Page 21

Search completed: June 20, 2004, 15:31:44  
Job time : 11842 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 06:33:06 ; Search time 1252 Seconds  
(without alignments)  
11277.090 Million cell updates/sec

Title: US-09-778-963A-1

Perfect score: 3082

Sequence: 1 GGCGTCGCGCGCGGCGGAGG.....aaaaaaaaaaaaaaaaa 3082

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10C\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3082	100.0	3082	9	US-09-778-963A-1
2	2949.4	95.7	2973	13	US-09-918-715-197
3	2806.4	91.1	2832	13	US-09-918-715-174
4	2806.4	91.1	2832	16	US-10-185-035-4
5	2500.4	81.1	11221	9	US-09-778-963A-3
6	2500.4	81.1	16000	16	US-10-185-035-11
7	1018.4	33.0	1021	16	US-10-321-039-13
8	734	23.8	753	13	US-10-027-632-25123
9	734	23.8	753	13	US-10-027-632-25124
10	734	23.8	753	13	US-10-027-632-25125
11	734	23.8	753	16	US-10-027-632-25123
12	734	23.8	753	16	US-10-027-632-25124
13	734	23.8	753	16	US-10-027-632-25125
14	694.4	22.5	696	15	US-10-029-386-20951

15	693.2	22.5	3020	13	US-09-918-715-292	Sequence 292, App
16	603.6	19.6	604	13	US-10-027-632-106883	Sequence 106883, A
17	603.6	19.6	604	13	US-10-027-632-106884	Sequence 106884, A
18	603.6	19.6	604	13	US-10-027-632-106885	Sequence 106885, A
19	603.6	19.6	604	16	US-10-027-632-106883	Sequence 106883, A
20	603.6	19.6	604	16	US-10-027-632-106884	Sequence 106884, A
21	603.6	19.6	604	16	US-10-027-632-106885	Sequence 106885, A
22	520	16.9	520	15	US-10-029-386-7251	Sequence 7251, App
23	519	16.8	523	16	US-10-185-035-12	Sequence 12, App
24	496	16.1	567	10	US-09-918-995-29627	Sequence 29627, A
25	425.4	13.8	469	10	US-09-918-995-2741	Sequence 2741, App
26	416	13.5	1744	15	US-10-197-666A-87	Sequence 87, App
27	283	9.2	368	9	US-09-864-761-21643	Sequence 21643, A
28	188	6.1	405	9	US-09-960-352-10273	Sequence 10273, A
29	151	4.9	597	10	US-09-873-546-4	Sequence 4, App
30	151	4.9	597	15	US-10-029-386-22936	Sequence 22936, A
31	127.4	4.1	1108	15	US-10-037-870-945	Sequence 945, App
32	127.4	4.1	1108	16	US-10-117-722-945	Sequence 945, App
33	127.4	4.1	4167	17	US-10-399-455-36	Sequence 36, App
34	119	3.9	212	13	US-10-085-783A-22243	Sequence 22243, A
35	119	3.9	212	16	US-10-242-535A-22243	Sequence 22243, A
36	118	3.8	2270	13	US-10-424-599-44604	Sequence 44604, A
37	117.4	3.8	327	9	US-09-960-352-4630	Sequence 4630, App
38	117.4	3.8	524	13	US-10-424-599-11691	Sequence 11691, A
39	117.2	3.8	378	13	US-10-424-599-95181	Sequence 95181, A
40	116.8	3.8	1416	16	US-10-295-027-688	Sequence 688, App
41	116.8	3.8	1494	14	US-10-062-254-259	Sequence 259, App
42	116.8	3.8	1844	10	US-09-814-353-19965	Sequence 19965, A
43	116.6	3.8	144	14	US-10-114-893-208	Sequence 208, App
44	116.6	3.8	223	13	US-10-085-783A-8175	Sequence 8175, App
45	116.6	3.8	223	16	US-10-242-535A-8175	Sequence 8175, App

## ALIGNMENTS

RESULT 1	
US-09-778-963A-1	
; Sequence 1, Application US/09778963A	
; Patent No. US20020115172A1	
GENERAL INFORMATION:	
; APPLICANT: NEELAM, Beena et al	
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,	
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE	
; FILE REFERENCE: CLO01112	
; CURRENT APPLICATION NUMBER: US/09/778, 963A	
; CURRENT FILING DATE: 2001-02-08	
; NUMBER OF SEQ ID NOS: 5	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 1	
; LENGTH: 3082	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-09-778-963A-1	
QY	Query Match 100.0%; Score 3082; DB 9; Length 3082;
QY	Best Local Similarity 100.0%; Pred. No. 0;
QY	Matches 3082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GGCGTCGCGCGCGGAGGAAGACCGCGCCGAGCCGCGCTCCGAGCAGCGCAGG 60
DB	1 GGCGTCGCGCGCGGAGGAAGACCGCGCCGAGCCGCGCTCCGAGCAGCGCAGG 60
QY	61 GAGATCCCGCGAGTAGACCGGAGGCACACAGACTTGGAGGCTCGGCGTGA 120
DB	61 GAGATCCCGCGAGTAGACCGGAGGCACACAGACTTGGAGGCTCGGCGTGA 120
QY	121 GAGCAGCAGCTCCCGCAGCTCCCGGCGCTTCAGCAGCTCTTGAGCCGTGCGA 180
DB	121 GAGCAGCAGCTCCCGCAGCTCCCGGCGCTTCAGCAGCTCTTGAGCCGTGCGA 180
QY	181 GGCGCGCGCGCGCATTCACACCGCGCAGCCGAGCAGCTTGTCCAGCGGAGCACTGC 240

Db	181	GGCCCGGCCCCGACATTTCCAGCCCCGAGGCCATGATGAAACCTTTGTCCAGCGGAACTGC	240
Qy	241	ACGCTCAGTGTGCCCGCCAAAACTCATACCGCATGCTGTGTGCTGTGGTCCCTTCGGGTG	300
Db	241	ACGCTCAGTGTGCCCGCCAAAAACTCATACCGCATGCTGTGTGCTGTGGTCCCTTCGGGTG	300
Qy	301	GGCAAGAGTCCATGTGTCTCGCTTCCCTCAATGGCCGCTTTGAGGACCAATACACACC	360
Db	301	GGCAAGAGTCCATGTGTCTCGCTTCCCTCAATGGCCGCTTTGAGGACCAATACACACC	360
Qy	361	ACCATGAGGACTTCCACCGTAAGGATACAACATCCGCGGACATGTACCAAGCTCGAC	420
Db	361	ACCATGAGGACTTCCACCGTAAGGATACAACATCCGCGGACATGTACCAAGCTCGAC	420
Qy	421	ATCTGGAATACCTCTGGCAACAACCCCTTCCCGCATGCGAGGCTGTCCATCCTCACA	480
Db	421	ATCTGGAATACCTCTGGCAACAACCCCTTCCCGCATGCGAGGCTGTCCATCCTCACA	480
Qy	481	GGGAGATGTTTATCCTCGGTGTTCAGCCGTAATACCGGAGGCTTCGATAGAGGTCAAG	540
Db	481	GGGAGATGTTTATCCTCGGTGTTCAGCCGTAATACCGGAGGCTTCGATAGAGGTCAAG	540
Qy	541	CGCCTTCAGAACCAATCTGAGGTCAAGTCTCTGTTAAGAACAGACCAAGAGAGCG	600
Db	541	CGCCTTCAGAACCAATCTGAGGTCAAGTCTCTGTTAAGAACAGACCAAGAGAGCG	600
Qy	601	GGGAGCTGCCCATGTGTCATGTGTGCAACAAGAACACACAGGAGGCTGTGTCCGCGAC	660
Db	601	GGGAGCTGCCCATGTGTCATGTGTGCAACAAGAACACACAGGAGGCTGTGTCCGCGAC	660
Qy	661	GTGCCACACACCGAGGCCGAGCTGTGTGTGGGCGACAGAACTCCGCTTACTTCCAG	720
Db	661	GTGCCACACACCGAGGCCGAGCTGTGTGTGGGCGACAGAACTCCGCTTACTTCCAG	720
Qy	721	GTGTGGGCTCAAGAAAGAACCAACGTGAGCAGAGTCTTACGTGCTCTTCAGCATGTGCC	780
Db	721	GTGTGGGCTCAAGAAAGAACCAACGTGAGCAGAGTCTTACGTGCTCTTCAGCATGTGCC	780
Qy	781	AAGCTGCACACGAGATGAGCCCCGCCCTGCATTCGCAAGATCTCCGTGACATACGGTGAC	840
Db	781	AAGCTGCACACGAGATGAGCCCCGCCCTGCATTCGCAAGATCTCCGTGACATACGGTGAC	840
Qy	841	GCTTTCACCCCAAGGCCCTTCTGCATGCGCGCGCTCAGAGATGAGCGCTTATGGCATG	900
Db	841	GCTTTCACCCCAAGGCCCTTCTGCATGCGCGCGCTCAGAGATGAGCGCTTATGGCATG	900
Qy	901	GTCGTGCCCTTGGCCCGGCCGCCCAAGCTCAACAGTGACTCAAGTCACTCAAGGCCAAG	960
Db	901	GTCGTGCCCTTGGCCCGGCCGCCCAAGCTCAACAGTGACTCAAGTCACTCAAGGCCAAG	960
Qy	961	GTCCTTGGGAAAGCCAGAGCCCGTGAAGAGAACAAATGACCATCAAGTGAAGCGGAT	1020
Db	961	GTCCTTGGGAAAGCCAGAGCCCGTGAAGAGAACAAATGACCATCAAGTGAAGCGGAT	1020
Qy	1021	GCTGGGGGGGGGCTTGGCCAGTGCCTTCAAGGAGGTGGCCCAAGTGCCTCAATGGCGCA	1080
Db	1021	GCTGGGGGGGGGCTTGGCCAGTGCCTTCAAGGAGGTGGCCCAAGTGCCTCAATGGCGCA	1080
Qy	1081	TCTCCCAACCAAGAGGCCCGCGACAGAGTCTTGTTCACAGACTTAAAGCACCAAGCTGAGG	1140
Db	1081	TCTCCCAACCAAGAGGCCCGCGACAGAGTCTTGTTCACAGACTTAAAGCACCAAGCTGAGG	1140
Qy	1141	CCCCCGAGGCTGGGCTCTCGCAACATTCGTGTGCTTCTCAAGCTTCTCTGATCGGCTT	1200
Db	1141	CCCCCGAGGCTGGGCTCTCGCAACATTCGTGTGCTTCTCAAGCTTCTCTGATCGGCTT	1200
Qy	1201	GTCCACAGCTCTGTGTGTGTTTCACTCTCTCTGTGGGAGAACACATCTTGACGCTTCAA	1260
Db	1201	GTCCACAGCTCTGTGTGTGTTTCACTCTCTCTGTGGGAGAACACATCTTGACGCTTCAA	1260
Qy	1261	GAGTTAGCAGAGACTCAAGTTACACTTCTCTCTGTGGGTTGAAAGAAATGTTGATGC	1320

Db	1261	GAGTTAGCAGAGACTCAAGTTAACTCTCTCTCTGCGGTTGGAAGAAATGTTGATGC	1320
Qy	1321	CAGAGGGGTGAGATTGCTGCGTATATGAGGCTCTCTG3GACAAGCTTCAGATGAAAA	1380
Db	1331	CAGAGGGGTGAGATTGCTGCGTATATGAGGCTCTCTG3GACAAGCTTCAGATGAAAA	1380
Qy	1381	GGACACAGAAAGCCAGATGAGAAAGGTCTCTCTCTG3CATTAACACCAGTTGGTT	1440
Db	1381	GGACACAGAAAGCCAGATGAGAAAGGTCTCTCTCTG3CATTAACACCAGTTGGTT	1440
Qy	1441	TGGGTGGCAGCTGGAGAACTTCTCTCCAGCCTTGCACTCTTACGCTGTGTTCACT	1500
Db	1441	TGGGTGGCAGCTGGAGAACTTCTCTCCAGCCTTGCACTCTTACGCTGTGTTCACT	1500
Qy	1501	GCCCTTGCACCCCTCCCAACCCCGACACACAAAGTTGGCCCCCAGCTGCGGCTGCAC	1560
Db	1501	GCCCTTGCACCCCTCCCAACCCCGACACACAAAGTTGGCCCCCAGCTGCGGCTGCAC	1560
Qy	1561	ATTAGAGCAGATGAGACTGTGTGTGAAAGGGGGCGTGAGCACACTCTTAGACACAGCCCA	1620
Db	1561	ATTAGAGCAGATGAGACTGTGTGTGAAAGGGGGCGTGAGCACACTCTTAGACACAGCCCA	1620
Qy	1621	COACTTAGACCAAGCCCACTCTTGACCGGGTTCCTAAGCTCCTCTCTAGTCCCTCC	1680
Db	1621	COACTTAGACCAAGCCCACTCTTGACCGGGTTCCTAAGCTCCTCTCTAGTCCCTCC	1680
Qy	1681	GCCGACAGTGTGTGCTTTGTTGTGAGGTGGACGTGTTTCGTGATGATATGATGATGAA	1740
Db	1681	GCCGACAGTGTGTGCTTTGTTGTGAGGTGGACGTGTTTCGTGATGATATGATGATGAA	1740
Qy	1741	ATGGAATCATTTGTACTGTAAAGCTTAGTACTCCCTCTTG3GACAGGCCCTCACCCAG	1800
Db	1741	ATGGAATCATTTGTACTGTAAAGCTTAGTACTCCCTCTTG3GACAGGCCCTCACCCAG	1800
Qy	1801	TTTCAGATCAAGGCTCCACCCGGGAGGCGCTTCTCTCTG3CTCCAAACAGGGTTCCG	1860
Db	1801	TTTCAGATCAAGGCTCCACCCGGGAGGCGCTTCTCTCTG3CTCCAAACAGGGTTCCG	1860
Qy	1861	TGGCCTGTTTGCAGCTAGACATTTAGCCTCCGCAATTGAGTCTCAAGGTTTACAGCAATT	1920
Db	1861	TGGCCTGTTTGCAGCTAGACATTTAGCCTCCGCAATTGAGTCTCAAGGTTTACAGCAATT	1920
Qy	1921	GCAACAAGTGTGGGTGGGCAAGCCAGACATGCTTTTTTTTAAATGCTCCCATTTACAGAG	1980
Db	1921	GCAACAAGTGTGGGTGGGCAAGCCAGACATGCTTTTTTTTAAATGCTCCCATTTACAGAG	1980
Qy	1981	GATACCAACCGAGACTCGGAGGGGACAAGATAGACAGAGGCCACCTTTGTCCCTAGC	2040
Db	1981	GATACCAACCGAGACTCGGAGGGGACAAGATAGACAGAGGCCACCTTTGTCCCTAGC	2040
Qy	2041	AAATTCAGGGTACAGCTCCACTTGAACCAAGGCTGCCCTCTACGTGTCTCTTCAAG	2100
Db	2041	AAATTCAGGGTACAGCTCCACTTGAACCAAGGCTGCCCTCTACGTGTCTCTTCAAG	2100
Qy	2101	CATTTATTAACACCTTACTGGGCTGTGGGTTCACTGTGCTCTAGAAAACCAAGGGGTC	2160
Db	2101	CATTTATTAACACCTTACTGGGCTGTGGGTTCACTGTGCTCTAGAAAACCAAGGGGTC	2160
Qy	2161	CCAGTCTCTGCGCTTGCCCGGCTCTGCTGCCCAACAATTCTTGACACAACAGCGGTGG	2220
Db	2161	CCAGTCTCTGCGCTTGCCCGGCTCTGCTGCCCAACAATTCTTGACACAACAGCGGTGG	2220
Qy	2221	GAGGCGGGGAGAGACACTGGGAAACCGAACTGAGCTTGGGAGGGATCCACAGAAAAGC	2280
Db	2221	GAGGCGGGGAGAGACACTGGGAAACCGAACTGAGCTTGGGAGGGATCCACAGAAAAGC	2280
Qy	2281	TCAGGGCGGGCTTCTCCTTGTGCCCAGGATTTGGGCTATGCTGTGGTACCAACATGTATC	2340
Db	2281	TCAGGGCGGGCTTCTCCTTGTGCCCAGGATTTGGGCTATGCTGTGGTACCAACATGTATC	2340
Qy	2341	AGGCAATGCTGGGTTTGAACCCATTAACAAAGGCCCTTGTCACTAGCTTTAACAGATA	2400
Db	2341	AGGCAATGCTGGGTTTGAACCCATTAACAAAGGCCCTTGTCACTAGCTTTAACAGATA	2400



```

QY 2401 TATTGATTTTAACTCTTAACAATTAAGTTTAAAGGCTTAAGGACCTTAGT 2460
Db 2401 TATTGATTTTAACTCTCTTAACAATTAAGTTTAAAGGCTTAAGGACCTTAGT 2460
QY 2461 GATCTTATATGGGCTCTTCTGAGGTTCAAGAGGGTAACTCTCTCCAGGTCAAC 2520
Db 2461 GATCTTATATGGGCTCTTCTGAGGTTCAAGAGGGTAACTCTCTCCAGGTCAAC 2520
QY 2521 AGCAAGTCTGGGTGGCAGAGCAAGTAAAGCTAGGCTAGGCTTCAATATCCAGATGT 2580
Db 2521 AGCAAGTCTGGGTGGCAGAGCAAGTAAAGCTAGGCTAGGCTTCAATATCCAGATGT 2580
QY 2581 GCTCCCTCTCTGAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2640
Db 2581 GCTCCCTCTCTGAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2640
QY 2641 ACCCTCTCCCTAGATGAGTCTGAGGCTCCCTGATATATGTTGATACCAATGCCATG 2700
Db 2641 ACCCTCTCCCTAGATGAGTCTGAGGCTCCCTGATATATGTTGATACCAATGCCATG 2700
QY 2701 TATATACAGTACACACAGATGTACACACAGATGTACACATGCTCCAGCCCAAGCTTGC 2760
Db 2701 TATATACAGTACACACAGATGTACACACAGATGTACACATGCTCCAGCCCAAGCTTGC 2760
QY 2761 ATACTGCACTGCAACCCAGCTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2820
Db 2761 ATACTGCACTGCAACCCAGCTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2820
QY 2821 CCAACCTGCTCTGTCCCTTCCCACTGCTGAGCTTCTGAGCAGACAGGTA 2880
Db 2821 CCAACCTGCTCTGTCCCTTCCCACTGCTGAGCTTCTGAGCAGACAGGTA 2880
QY 2881 CTTGGCTGCAACCGGTGTGTGCGCCCTCTCAACCCAGGCAAGCCCGCAATGATC 2940
Db 2881 CTTGGCTGCAACCGGTGTGTGCGCCCTCTCAACCCAGGCAAGCCCGCAATGATC 2940
QY 2941 TCCGCTGACTATCAATAAAGTGGTTTGTACAAAAAATTTTTTTTTTTTTTTTTTTT 3000
Db 2941 TCCGCTGACTATCAATAAAGTGGTTTGTACAAAAAATTTTTTTTTTTTTTTTTTTT 3000
QY 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
Db 3001 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3060
QY 3061 AAAAAAAAAAAAAAAAAAAAAAAAAA 3082
Db 3061 AAAAAAAAAAAAAAAAAAAAAAAAAA 3082

```

```

RESULT 2
US-09-918-715-197
; Sequence 197, Application US/09918715
; Publication No. US20030017157A1
GENERAL INFORMATION:
APPLICANT: Brad St. Croix
APPLICANT: Bert Vogelstein
APPLICANT: Kenneth Kinzler
TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
FILE REFERENCE: 1107.00134
CURRENT APPLICATION NUMBER: US/09/918,715
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/222,599
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/224,360
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/282,850
PRIOR FILING DATE: 2000-04-11
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 197
LENGTH: 2973
TYPE: DNA

```

```

! ORGANISM: Homo sapiens
US-09-918-715-197
Query Match 95.7%; Score 2949.4; DB 13; Length 2973;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 GCGCGCCAGCCGCGGCTCCCGAGCAGCGAGGAGATCCCGCGCAGTGAACCCGG 85
Db 23 GCGCGCCAGCCGCGGCTCCCGAGCAGCGAGGAGATCCCGCGCAGTGAACCCGG 82
QY 86 AGCCACACAGACTCTGGAGAGCTCGGCGCTGAGAGCAGGAGCTCCCGCAGCTCC 145
Db 83 AGCCACACAGACTCTGGAGAGCTCGGCGCTGAGAGCAGGAGCTCCCGCAGCTCC 142
QY 146 CGGCGCTTCAAGGAGCTCTCTGAGGCTGCGCAGAGGCCCCGGCCATTTCCAGCCCC 205
Db 143 CGGCGCTTCAAGGAGCTCTCTGAGGCTGCGCAGAGGCCCCGGCCATTTCCAGCCCC 202
QY 206 GAGCCATGATGAAGACTTGTCCAGGGGAACTGCAAGCTCAGTGTGCCGCAAAACT 265
Db 203 GAGCCATGATGAAGACTTGTCCAGGGGAACTGCAAGCTCAGTGTGCCGCAAAACT 262
QY 266 CATACCCGATGATGCTGCTGGGCTCTCTCGGCTGGGCAAGCTCCATGCTGCTGCT 325
Db 263 CATACCCGATGATGCTGCTGGGCTCTCTCGGCTGGGCAAGCTCCATGCTGCTGCT 322
QY 326 TCTCAATGCGCGCTTGAAGACCAATGACACACCAATCAGAGACTTCCACCTGAAG 385
Db 323 TCTCAATGCGCGCTTGAAGACCAATGACACACCAATCAGAGACTTCCACCTGAAG 382
QY 386 TATCAACATCCGGGGAGACATGACAGCTCGACATCTGGATACCTGGGCAACAC 445
Db 383 TATCAACATCCGGGGAGACATGACAGCTCGACATCTGGATACCTGGGCAACAC 442
QY 446 CTTTCCCGCATGCGAGGCTGCTCATCTCAAGGAGATGCTTATCTGTGATTTCA 505
Db 443 CTTTCCCGCATGCGAGGCTGCTCATCTCAAGGAGATGCTTATCTGTGATTTCA 502
QY 506 GCTGATTAACCGGAGTCTCTGATGAGTCAAGCGCTTCAAGAGCATCTGAGAG 565
Db 503 GCTGATTAACCGGAGTCTCTGATGAGTCAAGCGCTTCAAGAGCATCTGAGAG 562
QY 566 TCAAGTCTGCTGTAAGAACAGCAAGAGGCGGGAGCTGCCATGCTCATCTGTG 625
Db 563 TCAAGTCTGCTGTAAGAACAGCAAGAGGCGGGAGCTGCCATGCTCATCTGTG 622
QY 626 GCAACAGAAAGCAACGCGAGCTGCGCGCAGGTGCCAGCAGCAGGCGAGCTGC 685
Db 623 GCAACAGAAAGCAACGCGAGCTGCGCGCAGGTGCCAGGCTGCCAGGCGAGCTGC 682
QY 686 TGTGTGCGGCGAGAGAACTCCGCTTACTTGAAGTGTGCGCAAGAAACCAACG 745
Db 683 TGTGTGCGGCGAGAGAACTCCGCTTACTTGAAGTGTGCGCAAGAAACCAACG 742
QY 746 TGAACGATTTTCTAGTGTCTTTCAGCATGSCCAAGCTGCCACAGAGTGAAGCCCCG 805
Db 743 TGAACGATTTTCTAGTGTCTTTCAGCATGSCCAAGCTGCCACAGAGTGAAGCCCCG 802
QY 806 CCTGATGCGAAGATCTCCGTGAGTACAGCTTCCAGCCAGGCGCTTCTTGA 865
Db 803 CCTGATGCGAAGATCTCCGTGAGTACAGCTTCCAGCCAGGCGCTTCTTGA 862
QY 866 TGGCGCGGTCAAGAGATGAGCGCTTATGAGCATGCTGCGCTTGGCCGCGGCCCA 925
Db 863 TGGCGCGGTCAAGAGATGAGCGCTTATGAGCATGCTGCGCTTGGCCGCGGCCCA 922
QY 926 GGTCAACAGTAACTCAAGTACATCAAGGCAAGGCTTCCGGAAGGCCAGGCGGTG 985
Db 923 GGTCAACAGTAACTCAAGTACATCAAGGCAAGGCTTCCGGAAGGCCAGGCGGTG 982
QY 986 AGAGGAGCAAGTGAACATCAGTGAAGGAGATGCTGGGCGCGGCTTGGCCAGTGC 1045

```

Db 983 AGAGGAGACAAGTGCACCATCCAGTGAAGGAGATGCTGGGGCTTGGCCAGTGC 1042  
Qy 1046 TTCCAGGAGAGTGGCCCAAGATGCACTGTGGCATCTCCCAACGAGGCCCGGAGCA 1105  
Db 1043 TTCCAGGAGAGTGGCCCAAGATGCACTGTGGCATCTCCCAACGAGGCCCGGAGCA 1102  
Qy 1106 GTCTTGTTACAGACCTTAGGAGACGAGCTGAGAGGCCCGGAGGCTGGCCTCCGACAT 1165  
Db 1103 GTCTTGTTACAGACCTTAGGAGACGAGCTGAGAGGCCCGGAGGCTGGCCTCCGACAT 1162  
Qy 1166 TGGTCTGCTCTTCACAGCTTTCTGTAGTCCGCTGTGTCACAGCTCTTGTGTGTCTAT 1225  
Db 1163 TGGTCTGCTCTTCACAGCTTTCTGTAGTCCGCTGTGTCACAGCTCTTGTGTGTCTAT 1222  
Qy 1226 CTTCCTGTGGAGAGACATCTGTGACGCTTCAGAGTTAGGACAGACTCAAGTTACA 1285  
Db 1223 CTTCCTGTGGAGAGACATCTGTGACGCTTCAGAGTTAGGACAGACTCAAGTTACA 1282  
Qy 1286 CCTTCCTCTCTGGGGGTGAGAAATGTTGATGCCAGAGGGGTGAGATTGCTGTGCA 1345  
Db 1283 CCTTCCTCTCTGGGGGTGAGAAATGTTGATGCCAGAGGGGTGAGATTGCTGTGCA 1342  
Qy 1346 TATGAGACCTCTGAGACAGAGCTCAGATGAAAGACAGAAAGCCAGATGAGAAAG 1405  
Db 1343 TATGAGACCTCTGAGACAGAGCTCAGATGAAAGACAGAAAGCCAGATGAGAAAG 1402  
Qy 1406 GTCTCTCTCTCTGTGACATTAACCCAGCTTGGGTGGAGCTGGGAGAACTTCTC 1465  
Db 1403 GTCTCTCTCTCTGTGACATTAACCCAGCTTGGGTGGAGCTGGGAGAACTTCTC 1462  
Qy 1466 TCCAGACCCCTGCAACTTAAGCTGTGTTAGCTGAGCTCTGACACCCCTCCACCCCA 1525  
Db 1463 TCCAGACCCCTGCAACTTAAGCTGTGTTAGCTGAGCTCTGACACCCCTCCACCCCA 1522  
Qy 1526 GCACACACACAAATTGGCCCCAGCTGAGCTGACATTGAGCAGTGAATCTGTGTCTG 1585  
Db 1523 GCACACACACAAATTGGCCCCAGCTGAGCTGACATTGAGCAGTGAATCTGTGTCTG 1582  
Qy 1586 AAGGGGGGCTGGCCACACTCTTAAGACACGCCCCACATTGACAGGCCCACTCTCTG 1645  
Db 1583 AAGGGGGGCTGGCCACACTCTTAAGACACGCCCCACATTGACAGGCCCACTCTCTG 1642  
Qy 1646 ACCGGGCTCTCAGGCTCTCTCTAGAGTCCCTCCGCGCCGACAGTTGTCCTTGTGTG 1705  
Db 1643 ACCGGGCTCTCAGGCTCTCTCTAGAGTCCCTCCGCGCCGACAGTTGTCCTTGTGTG 1702  
Qy 1706 TTGACAGCTGTTTCTGTGATGATAGTAAATGGAATCAATTGACTGTAAAGC 1765  
Db 1703 TTGACAGCTGTTTCTGTGATGATAGTAAATGGAATCAATTGACTGTAAAGC 1762  
Qy 1766 CTAGTACCTCCCTCTGGCCAGGCCCTTACCCAGTTCCAGATCCAGGCTTCCACCCGG 1825  
Db 1763 CTAGTACCTCCCTCTGGCCAGGCCCTTACCCAGTTCCAGATCCAGGCTTCCACCCGG 1822  
Qy 1826 ACGCTCTCTCTCTGCTCCCAAGAGGTTTCCGTTGCTGTTGACAGCTGACATTGA 1885  
Db 1823 ACGCTCTCTCTCTGCTCCCAAGAGGTTTCCGTTGCTGTTGACAGCTGACATTGA 1882  
Qy 1886 CCTCCGCAATTGAGCTCCAGGTTTACAGACAAATTGACAGAGCTGGGGTGGAGCCA 1945  
Db 1883 CCTCCGCAATTGAGCTCCAGGTTTACAGACAAATTGACAGAGCTGGGGTGGAGCCA 1942  
Qy 1946 GAGCTGCTTTTAAATGCTCCCATTTTACAGAGATCCAGAGCTCCGAGAGGGAGC 2005  
Db 1943 GAGCTGCTTTTAAATGCTCCCATTTTACAGAGATCCAGAGCTCCGAGAGGGAGC 2002  
Qy 2006 ACGATGAGACAGAGCCCACTTTGCTCCCTAGCAAAATTCAAGGTTACAGCTCCACTAG 2065  
Db 2003 ACGATGAGACAGAGCCCACTTTGCTCCCTAGCAAAATTCAAGGTTACAGCTCCACTAG 2062  
Qy 2066 AACGAGGCTGCTCTACTGTGCTGCTTCTCAAGCAATTATTAAGCACTTCTGGGTGC 2125  
Db 2063 AACGAGGCTGCTCTACTGTGCTGCTTCTCAAGCAATTATTAAGCACTTCTGGGTGC 2122

Qy 2126 TGGGTTCACTGTGTCTTAAGAAACCAAGAGGTCCCAAGTCTCTGCTGCGGCCCT 2185  
Db 2123 TGGGTTCACTGTGTCTTAAGAAACCAAGAGGTCCCAAGTCTCTGCTGCGGCCCT 2182  
Qy 2186 GCTGCCCAACCACTTTGACACACAGAGGTGGGAGAGGGGAGAGAGAGCTGGAGC 2245  
Db 2183 GCTGCCCAACCACTTTGACACACAGAGGTGGGAGAGGGGAGAGAGAGCTGGAGC 2242  
Qy 2246 CAGAACTGAGCTTGGAGAGGATCCGACAGAAAGCTCAGGGCGGGCTTCTCTTGTG 2305  
Db 2243 CAGAACTGAGCTTGGAGAGGATCCGACAGAAAGCTCAGGGCGGGCTTCTCTTGTG 2302  
Qy 2306 CGGATTTGGGCTTATGCTGGGATACCAACATGATCTCAGGCAATGATGGTTTGAACCA 2365  
Db 2303 CGGATTTGGGCTTATGCTGGGATACCAACATGATCTCAGGCAATGATGGTTTGAACCA 2362  
Qy 2366 AACCAAGAGCCCTGTATCAGCTCTTAACAGATATTTGATTTAATCTCTTAA 2425  
Db 2363 AACCAAGAGCCCTGTATCAGCTCTTAACAGATATTTGATTTAATCTCTTAA 2422  
Qy 2426 CATATTGAAGTTTAAAGGCCCTTAAGAACTTAAGATCTTCTATTGGGCTTTCTGAG 2485  
Db 2423 CATATTGAAGTTTAAAGGCCCTTAAGAACTTAAGATCTTCTATTGGGCTTTCTGAG 2482  
Qy 2486 TTCAAGAGGGTAAATTAATTCTTCCAGGTCACACAGAACTGTGGGTGGAGAGCA 2545  
Db 2483 TTCAAGAGGGTAAATTAATTCTTCCAGGTCACACAGAACTGTGGGTGGAGAGCA 2542  
Qy 2546 AGCTAGCCCTGGGCACTTCAATACACAGATGTCCTCTCTTATGCTTGGGCC 2605  
Db 2543 AGCTAGCCCTGGGCACTTCAATACACAGATGTCCTCTCTTATGCTTGGGCC 2602  
Qy 2606 TGGGCTCTTCAAGGCTTTGAGACATCTTGTCTCAACCTCTCTCTAGATCAGTCTG 2665  
Db 2603 TGGGCTCTTCAAGGCTTTGAGACATCTTGTCTCAACCTCTCTCTAGATCAGTCTG 2662  
Qy 2666 GGGTCCCTGATATTTGTGTACACACATGCCATGATATTAACAGTACACACATGAC 2725  
Db 2663 GGGTCCCTGATATTTGTGTACACACATGCCATGATATTAACAGTACACACATGAC 2722  
Qy 2726 ACACAGATGATACATGCTCCAGGCCAGCTCTGATACCTGACCTGACCCAGCCTT 2785  
Db 2723 ACACAGATGATACATGCTCCAGGCCAGCTCTGATACCTGACCTGACCCAGCCTT 2782  
Qy 2786 GGGCTCTGCTGCTGTGTCTCAAGAGCAGCTCCAACTGCTGTCTGCTCTCTCC 2845  
Db 2783 GGGCTCTGCTGCTGTGTCTCAAGAGCAGCTCCAACTGCTGTCTGCTCTCTCC 2842  
Qy 2846 CACCCAGCTGAGGCTTCTGAGACAGACAGTACCTTGGCTGACACCGGTGTGGGCC 2905  
Db 2843 CACCCAGCTGAGGCTTCTGAGACAGACAGTACCTTGGCTGACACCGGTGTGGGCC 2902  
Qy 2906 GCTCTACCCAGGACAGGCCCGCCAGCATGATCTCGGTATACATTAATAAGTG 2965  
Db 2903 GCTCTACCCAGGACAGGCCCGCCAGCATGATCTCGGTATACATTAATAAGTG 2962  
Qy 2966 GGTGTTTACA 2976  
Db 2963 GGTGTTTACA 2973

RESULT 3  
US-09-918-715-174  
; Sequence 174, Application US/09918715  
; Publication No. US20030017157A1  
; GENERAL INFORMATION:  
; APPLICANT: Brad St. Croix  
; APPLICANT: Bert Vogelstein  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107. 00134  
; CURRENT APPLICATION NUMBER: US/09/918. 715



```

QY 2119 TGGTGTGCTGGGTTCACTGTGTCTCTAGAAACCAAGGGTCCCACTCTGAGCTTGGC 2178
    |||||
Db 1945 TGGGTGTGCTGGTTCACCTGTGTCTCTAGAAACCAAGGGTCCCACTCTGAGCTTGGC 2004
QY 2179 CGCCCCGTGTCGCCCACTCTCTGACACACAGCGGTGGGGAGGCGGGAGAGCAGC 2238
    |||||
Db 2005 CGCCCCGTGTCGCCCACTCTCTGACACACAGCGGTGGGGAGGCGGGAGAGCAGC 2064
QY 2239 TGGGACCCAGAACTGAGCTGGGAGGAGTCCAGAGAAAAGCTCAGGCGGGTCTTCC 2298
    |||||
Db 2065 TGGGACCCAGAACTGAGCTGGGAGGAGTCCAGAGAAAAGCTCAGGCGGGTCTTCC 2124
QY 2299 TTGTGCTCGGGAGTTGGGCTATGTGGGTACACCATGTACTAGGAGATGGTGGTTTGA 2358
    |||||
Db 2125 TTGTGCTCGGGAGTTGGGCTATGTGGGTACACCATGTACTAGGAGATGGTGGTTTGA 2184
QY 2359 ACCCATAAACCAAGGCGCTTGTCTCATGAGCTCTTAAAGATATTTGTATTTTAACT 2418
    |||||
Db 2185 ACCCATAAACCAAGGCGCTTGTCTCATGAGCTCTTAAAGATATTTGTATTTTAACT 2244
QY 2419 CTCTAAACATATTGAAGTTTGAAGGCGCTTAAAGAACCTTAAGTATCTTATTTGGGCTT 2478
    |||||
Db 2245 CTCTAAACATATTGAAGTTTGAAGGCGCTTAAAGAACCTTAAGTATCTTATTTGGGCTT 2304
QY 2479 TCTGAGGTTCAAGAGGGGTAAGTAATCTCTCAGGTCACAGAGAGTGTGGTGGC 2538
    |||||
Db 2305 TCTGAGGTTCAAGAGGGGTAAGTAATCTCTCAGGTCACAGAGAGTGTGGTGGC 2364
QY 2539 AGAAGCAAGCTATGCGCTGGGCACTTCAATACCAAGATGTCTCTCTCTTGTATGCT 2598
    |||||
Db 2365 AGAAGCAAGCTATGCGCTGGGCACTTCAATACCAAGATGTCTCTCTCTTGTATGCT 2424
QY 2599 TGGCCCTTGGGGCTTCAAGGCTTTGGGACATCTTGTCTTAAACCTTCTCCTTAAGTACG 2658
    |||||
Db 2425 TGGCCCTTGGGGCTTCAAGGCTTTGGGACATCTTGTCTTAAACCTTCTCCTTAAGTACG 2484
QY 2659 TCTGTAGAGGCTCCCTGTATATTTGTGTACCAATGCCCCCATGATATTAAGATACACAA 2718
    |||||
Db 2485 TCTGTAGAGGCTCCCTGTATATTTGTGTACCAATGCCCCCATGATATTAAGATACACAA 2544
QY 2719 GATGTACACACAGATGTATACATGTCTCAGGCCCAAGCTTGTGATACCTGACCTGACACC 2778
    |||||
Db 2545 GATGTACACACAGATGTATACATGTCTCAGGCCCAAGCTTGTGATACCTGACCTGACACC 2604
QY 2779 CAGCCTTGGCGCTCTGCTGCTGTGTCTCAAGAGACAGCTCAACCTGTGCTCTGTCC 2838
    |||||
Db 2605 CAGCCTTGGCGCTCTGCTGCTGTGTCTCAAGAGACAGCTCAACCTGTGCTCTGTCC 2664
QY 2839 CTTTCCCAACCCACATGCTGAGCTTCTGTAGACAGACAGGTACTTGGCTGACCGGTGT 2898
    |||||
Db 2665 CTTTCCCAACCCACATGCTGAGCTTCTGTAGACAGACAGGTACTTGGCTGACCGGTGT 2724
QY 2899 GTGGCCCGCTCTCACCAGGCAACAGCCCGCCACCATGTGATCTCGGTGTACATATCAAT 2958
    |||||
Db 2725 GTGGCCCGCTCTCACCAGGCAACAGCCCGCCACCATGTGATCTCGGTGTACATATCAAT 2784
QY 2959 AAAAGTGGGTTGTATACAAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3006
    |||||
Db 2785 AAAAGTGGGTTGTATACAAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2832

```

## RESULT 4

```

US-10-185-035-4
; Sequence 4, Application US/10185035
; Publication No. US20040005706A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF DEXRASI EXPRESSION
; FILE REFERENCE: RTS-0377
; CURRENT APPLICATION NUMBER: US/10/185,035
; CURRENT FILING DATE: 2002-06-28

```

```

; NUMBER OF SEQ ID NOS: 135
; SEQ ID NO 4
; LENGTH: 2832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(837)
US-10-185-035-4

```

```

Query Match          91.1%; Score 2806.4; DB 16; Length 2832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 199 CAGCCCCAGCCATGATGAAGCTTTGTCCAGCGGGAACTGACGCTCAGTGTGCGGC 258
    |||||
Db 25 CAGCCCCAGCCATGATGAAGCTTTGTCCAGCGGGAACTGACGCTCAGTGTGCGGC 84
QY 259 AAAAAGCTATCCGATGTGGTGTGGTGTCTCTCGGGTGGGCAAGAGCTCCATCGTG 318
    |||||
Db 85 AAAAAGCTATCCGATGTGGTGTGGTGTCTCTCGGGTGGGCAAGAGCTCCATCGTG 144
QY 319 TCTCGCTTCTCAATGCGCGCTTTGAGGACCAATACACCCATCAGAGGACTTCCAC 378
    |||||
Db 145 TCTCGCTTCTCAATGCGCGCTTTGAGGACCAATACACCCATCAGAGGACTTCCAC 204
QY 379 CGTAAAGTATACACATCCGCGGCGCATGTATCCAGCTCGACATCTGATATCTTGGC 438
    |||||
Db 205 CGTAAAGTATACACATCCGCGGCGCATGTATCCAGCTCGACATCTGATATCTTGGC 264
QY 439 AACCAACCCCTTCCCGCCATGCGGAGGCTGTTCATCTCCAGAGGGAGTCTTCACTCG 498
    |||||
Db 265 AACCAACCCCTTCCCGCCATGCGGAGGCTGTTCATCTCCAGAGGGAGTCTTCACTCG 324
QY 499 GTTTCAGCTGATATTAACCGGAGTCTTTCATGAGTCAAGCGCTTCAAGAGCAGATC 558
    |||||
Db 325 GTTTCAGCTGATATTAACCGGAGTCTTTCATGAGTCAAGCGCTTCAAGAGCAGATC 384
QY 559 CTGAGAGTCAAGTCTGCTTGAAGAACCAAGAGGCGGGAGCTTCCATGTGTC 618
    |||||
Db 385 CTGAGAGTCAAGTCTGCTTGAAGAACCAAGAGGCGGGAGCTTCCATGTGTC 444
QY 619 ATCTGTGCAACAAGAAAGCAACAGCGGAGCTGTGCGCCAGGTGCCCCACCAAGAGGCC 678
    |||||
Db 445 ATCTGTGCAACAAGAAAGCAACAGCGGAGCTGTGCGCCAGGTGCCCCACCAAGAGGCC 504
QY 679 GAGCTGTGTGTGTGGGCGAGAGAACTCGCCTTACTTGAAGTGTGTGCGCCAAAGAAAC 738
    |||||
Db 505 GAGCTGTGTGTGTGGGCGAGAGAACTCGCCTTACTTGAAGTGTGTGCGCCAAAGAAAC 564
QY 739 ACCAAGTGTGAAGATGTGTCTAGTGTCTTCAAGATGAGCCAGCTGCCACAGAGATG 798
    |||||
Db 565 ACCAAGTGTGAAGATGTGTCTAGTGTCTTCAAGATGAGCCAGCTGCCACAGAGATG 624
QY 799 AGCCCCCGCTTCAATGCAAGATCTCCGTGAGTACGCTTCAAGCTTCAAGCCAGAGGCC 858
    |||||
Db 625 AGCCCCCGCTTCAATGCAAGATCTCCGTGAGTACGCTTCAAGCTTCAAGCCAGAGGCC 684
QY 859 TTCTGATGCGCGCGGTCAAGAGATGAGCGCTTATGAGTGTGTGTGCGCCCTTGGCGGC 918
    |||||
Db 685 TTCTGATGCGCGCGGTCAAGAGATGAGCGCTTATGAGTGTGTGTGCGCCCTTGGCGGC 744
QY 919 CGCCCCAGCGTCAACAGTACTCAAGTACTCAAGGCAAGGCTCTTGGGAGAGGCGAG 978
    |||||
Db 745 CGCCCCAGCGTCAACAGTACTCAAGTACTCAAGGCAAGGCTCTTGGGAGAGGCGAG 804
QY 979 GCCCGTGAAGAGGCAAGAGTCAACATCCAGTGAACGAGGATGTGTGGGCGGGGCTTGGC 1038
    |||||
Db 805 GCCCGTGAAGAGGCAAGAGTCAACATCCAGTGAACGAGGATGTGTGGGCGGGGCTTGGC 864
QY 1039 CAGTGTCTTCAAGGAGTGTGCGCCAGATGCCACTGTGCGGATCTTCCACCGAGGCCCC 1098
    |||||
Db 865 CAGTGTCTTCAAGGAGTGTGCGCCAGATGCCACTGTGCGGATCTTCCACCGAGGCCCC 924

```



TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-778-963a-3

Query Match 81.1%; Score 2500.4; DB 9; Length 11221;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2504; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 469 TCCATCTCAACAGGAGATGCTTCATCTCTGTTGTTCAAGCTGGATTAACGGGAGTCTTTC 528
DB 7681 TCTCTCTCGACAGGAGATGCTTCATCTCTGTTGTTCAAGCTGGATTAACGGGAGTCTTTC 7740
QY 529 GATGAGTCAAGAGGCTTTCAGAGCAGATCTGAGGTCAAGTCTGCTGTAAGAACAG 588
DB 7741 GATGAGTCAAGAGGCTTTCAGAGCAGATCTGAGGTCAAGTCTGCTGTAAGAACAG 7800
QY 589 ACCAAGAGAGCGCGAGCTGCTCCATGCTATCTGTGGACAAGAACAGACACGCGCAG 648
DB 7801 ACCAAGAGAGCGCGAGCTGCTCCATGCTATCTGTGGACAAGAACAGACACGCGCAG 7860
QY 649 CTGTGCGCGAGGTGCGCACACACCGAGGCGAGCTGTGTGCGGCGACAGAACTCC 708
DB 7861 CTGTGCGCGAGGTGCGCACACCGAGGCGAGCTGTGTGCGGCGACAGAACTCC 7920
QY 709 GCTTACTTGAAGTGTGTGCGCAAGAAACACCAAGTGAAGAGATGTTTACGTGCTC 768
DB 7921 GCTTACTTGAAGTGTGTGCGCAAGAAACACCAAGTGAAGATGTTTACGTGCTC 7980
QY 769 TTTCAGATGAGCCAGCTGTCACAGAGATGAGCCCGCTGATGCAAGATCTTCCTG 828
DB 7981 TTTCAGATGAGCCAGCTGTCACAGAGATGAGCCCGCTGATGCAAGATCTTCCTG 8040
QY 829 CAGTACGATGACGCTTTCACACCGAGGCTTCTGATGAGCGCGCTCAAGAGATGAGC 888
DB 8041 CAGTACGATGAGGCTTTCACACCGAGGCTTCTGATGAGCGCGCTCAAGAGATGAGC 8100
QY 889 GCTTATGAGTGTGTGTGCGCTTGTGCGCGCGCGCGAGCTCAAGTGAAGTCAAGTAC 948
DB 8101 GCTTATGAGTGTGTGTGCGCTTGTGCGCGCGCGCGAGCTCAAGTGAAGTCAAGTAC 8160
QY 949 ATTCAGAGCCAGAGTCTTTCAGAGAGGCGCGCTGAGAGAGACAGTGCACATCCAG 1008
DB 8161 ATTCAGAGCCAGAGTCTTTCAGAGAGGCGCGCTGAGAGAGACAGTGCACATCCAG 8220
QY 1009 TGAGGAGGAGATGCTGAGGCGGCGCTTGTGCGAGTGTGCTGAGAGAGTGGCCCAAGTAC 1068
DB 8221 TGAGGAGGAGATGCTGAGGCGGCGCTTGTGCGAGTGTGCTGAGAGAGTGGCCCAAGTAC 8280
QY 1069 CCACTGTGCGATCTTCCACAGAGGCGCGCGAGAGTCTTGTTCACAGACCTTAGGCA 1128
DB 8281 CCACTGTGCGATCTTCCACAGAGGCGCGCGAGAGTCTTGTTCACAGACCTTAGGCA 8340
QY 1129 CAGAGCTGAGGCGCGCGCGCGCTGCGCTGCGCAATTCGTGCTTCTTCACAGCTTTC 1188
DB 8341 CAGAGCTGAGGCGCGCGCGCGCTGCGCTGCGCAATTCGTGCTTCTTCACAGCTTTC 8400
QY 1189 CTGAGTCCGCTTGTTCACAGCTCTTGTGTTTATCTCTCTGAGGAGACATCT 1248
DB 8401 CTGAGTCCGCTTGTTCACAGCTCTTGTGTTTATCTCTCTGAGGAGACATCT 8460
QY 1249 CTGAGCTCTCAAGATGAGGAGACCTCAAGTTCACCTTCTCTCTGAGGTTGAG 1308
DB 8461 CTGAGCTCTCAAGATGAGGAGACCTCAAGTTCACCTTCTCTCTGAGGTTGAG 8520
QY 1309 AAATGTATGCGACAGAGGAGTGAAGATGCTGCGTATATGAGAGCTCTGAGAACCC 1368
DB 8521 AAATGTATGCGACAGAGGAGTGAAGATGCTGCGTATATGAGAGCTCTGAGAACCC 8580
QY 1369 TCAGATGAGAGAGACAGAGAGGCGAGATGAGAAAGTCTCTCTCTGAGTAACA 1428
DB 8581 TCAGATGAGAGAGACAGAGAGGCGAGATGAGAAAGTCTCTCTCTGAGTAACA 8640
QY 1429 CCCAGCTGTTGTGGGTGAGCTGAGAGAACTTCTCTCCAGCCCTGCAACTTACGC 1488
```

```
DB 8641 CCAGCTGTTGTGGGTGAGCTGAGAGAACTTCTCTCCAGCCCTGCAACTTACGC 8700
QY 1489 TCTGTTCAGTGCCTCTGACACCCCTCCACCCCGACACACACAGAGTTGGCCCCA 1548
DB 8701 TCTGTTCAGTGCCTCTGACACCCCTCCACCCCGACACACAGAGTTGGCCCCA 8760
QY 1549 GCTGGCGCTGACATTGAGCAGTGAAGCTGTGTCTGAAGGGGGGCTGAGCACCTCT 1608
DB 8761 GCTGGCGCTGACATTGAGCAGTGAAGCTGTGTCTGAAGGGGGGCTGAGCACCTCT 8820
QY 1609 AGACACGCGCCACCTTAGACACAGCCACCTCTGACCGGTTCCAGCTCTCTC 1668
DB 8821 AGACACGCGCCACCTTAGACACAGCCACCTCTGACCGGTTCTCAGCTCTCTC 8880
QY 1669 CTAGGTCCCTCCCGCAGATGTGCTTGTGTGTGAGAGCTGTTTGTGTGATGT 1728
DB 8881 CTAGGTCCCTCCCGCAGATGTGCTTGTGTGTGAGAGCTGTTTGTGTGATGT 8940
QY 1729 ATAGTATGAGAAATGGAATCATTTGATCTGTAAGCTTAGTACTCCTCTGAGCAG 1788
DB 8941 ATAGTATGAGAAATGGAATCATTTGATCTGTAAGCTTAGTACTCCTCTGAGCAG 9000
QY 1789 GGCCTCACCCAGTTCAATTCACAGGCTCTCAACCGGAGAGGCTTCTCTGCTCCCA 1848
DB 9001 GGCCTCACCCAGTTCAATTCACAGGCTCTCAACCGGAGAGGCTTCTCTGCTCCCA 9060
QY 1849 ACAGGTTTCCGTGCGCTGTTTGAAGTATGATTCATTCAGCTTCCAGCTCCACGCT 1908
DB 9061 ACAGGTTTCCGTGCGCTGTTTGAAGTATGATTCATTCAGCTTCCAGCTCCACGCT 9120
QY 1909 TTACAGCAATTTGACAGAGCGTGGGTGGAGGCGAGAGCTGTTTATATGCTCC 1968
DB 9121 TTACAGCAATTTGACAGAGCGTGGGTGGAGGCGAGAGCTGTTTATATGCTCC 9180
QY 1969 CATTTCAGAGAGATACACCGAGCTGAGAGGAGAGACAGATGAGACACAGGCCCCA 2028
DB 9181 CATTTCAGAGAGATACACCGAGCTGAGAGGAGAGACAGATGAGACACAGGCCCCA 9240
QY 2029 TTGTCCCTGACAAATTCAGAGTCAAGCTCCAGCTTGAACAGAGCTGCTTACTGTGC 2088
DB 9241 TTGTCCCTGACAAATTCAGAGTCAAGCTCCAGCTTGAACAGAGCTGCTTACTGTGC 9300
QY 2089 TCGTTCCTCAAGCAATTTATTAAGACCTACTGGGTGCTGAGTTCATGTCCTAGAGAA 2148
DB 9301 TCGTTCCTCAAGCAATTTATTAAGACCTACTGGGTGCTGAGTTCATGTCCTAGAGAA 9360
QY 2149 CCAAGAGGTCCTCAGTCTGCTGCTGCGCGCGCTGCTGCTCCACACACCTTCTGACA 2208
DB 9361 CCAAGAGGTCCTCAGTCTGCTGCTGCGCGCGCTGCTGCTCCACACACCTTCTGACA 9420
QY 2209 CACAGCGTGGGAGAGGCGGAGAGAGACAGTGGGACCCAGAACTGAGCTGGAGATC 2268
DB 9421 CACAGCGTGGGAGAGGCGGAGAGAGACAGTGGGACCCAGAACTGAGCTGGAGATC 9480
QY 2269 CAGAGAGAACTCAGAGGCGGCTTCTCTCTGCGCGGAGTTGGGCTATGCGGATC 2328
DB 9481 CAGAGAGAACTCAGAGGCGGCTTCTCTCTGCGCGGAGTTGGGCTATGCGGATC 9540
QY 2329 CACATGTACTCAGAGCTGTGGTGTGAAACCCATTAACCAAGGCGCTGTCTATCAGC 2388
DB 9541 CACATGTACTCAGAGCTGTGGTGTGAAACCCATTAACCAAGGCGCTGTCTATCAGC 9600
QY 2389 TCTTAAACAAGTAATTTTGTATTTTATCTCTCTTAAACATATGAAATTTTAAAGGCTCTA 2448
DB 9601 TCTTAAACAAGTAATTTTGTATTTTATCTCTCTTAAACATATGAAATTTTAAAGGCTCTA 9660
QY 2449 AGAACCCTAGTATCTTCTATTTGGTCTTCTCTGAGGTTCAAGAGGAGTAACTTCC 2508
DB 9661 AGAACCCTAGTATCTTCTATTTGGTCTTCTCTGAGGTTCAAGAGGAGTAACTTCC 9720
QY 2509 TCAGGTCAACAGCAAGTCTGTGGGTGAGAGAAAGTAAAGCTGAGCTTCACTTAC 2568
```



Db 9721 TCACAGTACACAGCAGTCTGTGGTGGCAGAAAGCAAGCTAGCGCTGGGATTCAGTAC 9780  
 Qy 2569 ATACACAGATGTGCTCCCTCTTGTGATGCTGGGCGCCCTTCAAGGCTTTGGAC 2628  
 Db 9781 ATACACAGATGTGCTCCCTCTTGTGATGCTGGGCGCCCTTCAAGGCTTTGGAC 9840  
 Qy 2629 ATCTGTGCTCAACCCCTCTTGTGATGCTGGGCGCCCTTCAAGGCTTTGGAC 2688  
 Db 9841 ATCTGTGCTCAACCCCTCTTGTGATGCTGGGCGCCCTTCAAGGCTTTGGAC 9900  
 Qy 2689 ACCATGCCATGTATATACAGTACACAGATGTACACAGATGTACACAGTGTCCAG 2748  
 Db 9901 ACCATGCCATGTATATACAGTACACAGATGTACACAGATGTACACAGTGTCCAG 9960  
 Qy 2749 CCCCAGCTCTGATACCTGACCTGACCCCAAGCTTGGCCCTTGGCTGTGTGCTC 2808  
 Db 9961 CCCCAGCTCTGATACCTGACCTGACCCCAAGCTTGGCCCTTGGCTGTGTGCTC 10020  
 Qy 2809 AAGCAGAGAGCTCCAAACCTGCTGTGCTTCCCAACCAAGCTGAGCTTGTGA 2868  
 Db 10021 AAGCAGAGAGCTCCAAACCTGCTGTGCTTCCCAACCAAGCTGAGCTTGTGA 10080  
 Qy 2869 GCAGACAGAGTACCTTGGCTGACCGGCTGTGTGCTTCCCAAGCAGCAGCCCG 2928  
 Db 10081 GCAGACAGAGTACCTTGGCTGACCGGCTGTGTGCTTCCCAAGCAGCAGCCCG 10140  
 Qy 2929 CCACCATGATCTCGGTGTACATCTATATAAAGTGGGTTGTACAA 2978  
 Db 10141 CCACCATGATCTCGGTGTACATCTATATAAAGTGGGTTGTACAA 10190

## RESULT 6

US-10-185-035-11  
 ; Sequence 11, Application US/10185035  
 ; Publication No. US20040005706a1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brett P. Monia  
 ; APPLICANT: Kenneth W. Dobie  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF DEXRASI EXPRESSION  
 ; FILE REFERENCE: RUS-0377  
 ; CURRENT APPLICATION NUMBER: US/10/185,035  
 ; CURRENT FILING DATE: 2002-06-28  
 ; NUMBER OF SEQ ID NOS: 135  
 ; SEQ ID NO 11  
 ; LENGTH: 16000  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-185-035-11

Query Match 81.1%; Score 2500.4; DB 16; Length 16000;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2504; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 469 TCATCTCCACAGGAGATGTCTTCACTGTGTGTCAAGCTGATTAACCGGAGTCTTC 528  
 Db 11913 TCTCTCCCTGAGGGAGATGTCTTCACTGTGTGTCAAGCTGATTAACCGGAGTCTTC 11972  
 Qy 529 GATGAGTCAAGCGCTTCAAGAGAGATCTGTGAGTCAAGTCTGCTGGAAGCAAG 588  
 Db 11973 GATGAGTCAAGCGCTTCAAGAGAGATCTGTGAGTCAAGTCTGCTGGAAGCAAG 12032  
 Qy 589 ACCAAGAGAGGCGGCGAGAGCTGCCATGTGATCTGTGCAACAAAGAGCAACCGCGAG 648  
 Db 12033 ACCAAGAGAGGCGGCGAGAGCTGCCATGTGATCTGTGCAACAAAGAGCAACCGCGAG 12092  
 Qy 649 CTGTGCGCCAGAGTCCCAACCAACCGAGCTGTGTGTGCGGCGACGAGAATCTC 708  
 Db 12093 CTGTGCGCCAGAGTCCCAACCAACCGAGCTGTGTGTGCGGCGACGAGAATCTC 12152  
 Qy 709 GCTTACTTGAGGTGTGCGCCAAAGAAACACCAAGTGTGAGAGATGTTCTAGGTGCTC 768  
 Db 12153 GCTTACTTGAGGTGTGCGCCAAAGAAACACCAAGTGTGAGAGATGTTCTAGGTGCTC 12212

Qy 769 TTCAGATGCGCAAGCTGCCACAGATGAGCCCGCTGTGATTCGACAGATCTCCGTG 828  
 Db 12213 TTCAGATGCGCAAGCTGCCACAGATGAGCCCGCTGTGATTCGACAGATCTCCGTG 12272  
 Qy 829 CAGTACGCTGAGCGCTTTCACCCCAAGGCTTCTGTGATGCGCGGTCAAGAGATGAC 888  
 Db 12273 CAGTACGCTGAGCGCTTTCACCCCAAGGCTTCTGTGATGCGCGGTCAAGAGATGAC 12332  
 Qy 889 GCTTATGAGATGTGCTGCGCTTTCAGCGCGCGCCCGAGGCTCAAGAGATCTCAAGTAC 948  
 Db 12333 GCTTATGAGATGTGCTGCGCTTTCAGCGCGCGCCCGAGGCTCAAGAGATCTCAAGTAC 12392  
 Qy 949 ATCAAGGCCAAGTCTTTCGGAAGGCCAGGCGGTGAGAGAGCAAGTGTACATCCAG 1008  
 Db 12393 ATCAAGGCCAAGTCTTTCGGAAGGCCAGGCGGTGAGAGAGCAAGTGTACATCCAG 12452  
 Qy 1009 TGAAGGAGAGTGTGCGGCGGAGCTTGGCCAGTGTCTTCAAGGAGTGTGCTCCAGATGC 1068  
 Db 12453 TGAAGGAGAGTGTGCGGCGGAGCTTGGCCAGTGTCTTCAAGGAGTGTGCTCCAGATGC 12512  
 Qy 1069 CCACTGTGCGATCTCCCAACCGAGGCGCCGAGAGAGCTTGTTCACAGACCTTAGGCA 1128  
 Db 12513 CCACTGTGCGATCTCCCAACCGAGGCGCCGAGAGAGCTTGTTCACAGACCTTAGGCA 12572  
 Qy 1129 CCAAGCTGAGAGCGCCCGGCGGCTGAGCTCCGACATTCGTGTGCTTCTCAAGCTTTC 1188  
 Db 12573 CCAAGCTGAGAGCGCCCGGCGGCTGAGCTCCGACATTCGTGTGCTTCTCAAGCTTTC 12632  
 Qy 1189 CTGAGTCCGCTTGTTCACAGCTCTTGTGTGATTCATCTCTGTGTGAGAGACATCT 1248  
 Db 12633 CTGAGTCCGCTTGTTCACAGCTCTTGTGTGATTCATCTCTGTGTGAGAGACATCT 12692  
 Qy 1249 CTGAGCTCTCAAGATTAAGGAGAGATCTCAAGTAAACCTTCTCTGTGGGTTGGAG 1308  
 Db 12693 CTGAGCTCTCAAGATTAAGGAGAGATCTCAAGTAAACCTTCTCTGTGGGTTGGAG 12752  
 Qy 1309 AAATGTTATGCTCAGAGGGGTGAGGATTTGCTGCGCATATGAGAGCTCTCTGAGCAAGCC 1368  
 Db 12753 AAATGTTATGCTCAGAGGGGTGAGGATTTGCTGCGCATATGAGAGCTCTCTGAGCAAGCC 12812  
 Qy 1369 TCAGATGAAAAAGACACAGAGGCGCAGATGAGAAAGTCTCTCTCTGTGGCATTAACA 1428  
 Db 12813 TCAGATGAAAAAGACACAGAGGCGCAGATGAGAAAGTCTCTCTCTGTGGCATTAACA 12872  
 Qy 1429 CCCAGCTGTTGGTGGTGGAGCTGAGGAACTTCTCTCCAGCGCTTCAACTCTTAAGC 1488  
 Db 12873 CCCAGCTGTTGGTGGTGGAGCTGAGGAACTTCTCTCCAGCGCTTCAACTCTTAAGC 12932  
 Qy 1489 TCTGTTGAGCTGAGCTGTGACACCCGCTCCAGCCCGCAGACACACAAGTTGGCGCCA 1548  
 Db 12933 TCTGTTGAGCTGAGCTGTGACACCCGCTCCAGCCCGCAGACACACAAGTTGGCGCCA 12992  
 Qy 1549 GCTGCGCTGACATTTAGCCAGTGAATCTGTGTCTGAAAGGGGGGTGGGCAACATCTCT 1608  
 Db 12993 GCTGCGCTGACATTTAGCCAGTGAATCTGTGTCTGAAAGGGGGGTGGGCAACATCTCT 13052  
 Qy 1609 AGACCAAGCCCAACCACTTAAACACAGGCCCACTCTGACCGGCTTCTCAAGCTCTCTC 1668  
 Db 13053 AGACCAAGCCCAACCACTTAAACACAGGCCCACTCTGACCGGCTTCTCAAGCTCTCTC 13112  
 Qy 1669 CTAGGTCCCTCGCGCCGACAGTTGAGCTTGTGTGAGTGTGAGAGCTGTTTGTGTGATGT 1728  
 Db 13113 CTAGGTCCCTCGCGCCGACAGTTGAGCTTGTGTGAGTGTGAGAGCTGTTTGTGTGATGT 13172  
 Qy 1729 ATAGTAGTAAGTAAGTAATCATTTGTAAGTGAAGCTCTGATGATCTCTCTTGGCCAG 1788  
 Db 13173 ATAGTAGTAAGTAAGTAATCATTTGTAAGTGAAGCTCTGATGATCTCTCTTGGCCAG 13232  
 Qy 1789 GCGCTCACCCAGTTCAATTCACAGGCTCTCACCCGAGAGCGCTTCTCTCTGCTCCAA 1848  
 Db 13233 GCGCTCACCCAGTTCAATTCACAGGCTCTCACCCGAGAGCGCTTCTCTCTGCTCCAA 13292  
 Qy 1849 ACAGGTTTCGTGGCTGTTTGAAGTGAATGACCTCCGCAATGAGCTCCAGCGT 1908

Db	13293	ACAGGGTTCCGTGGCCGTTTGGACGGTAGACATTGACCTCCGCAATGAGCTCCAGGGT	13355
QY	1909	TTACGACAAATTGCACAAGCGCTGGGGTGGGCAAGCCAGACCTGCTTTTATATGCTCC	1968
Db	13353	TTACGACAAATTGCACAAGCGCTGGGGTGGGCAAGCCAGACCTGCTTTTATATGCTCC	13412
QY	1969	CATTTCACAGAGATACCAACGGAACCTCGAGGGGACACAGATGAGCACAGAGCCCACTT	2028
Db	13413	CATTTCACAGAGATACCAACGGAACCTCGAGGGGACACAGATGAGCACAGAGCCCACTT	13472
QY	2029	TTGTCCCTAGCAAAATTCAGGGTACAGCTCCACCTAGAACAGAGCTGCCCTTACTGTGC	2088
Db	13473	TTGTCCCTAGCAAAATTCAGGGTACAGCTCCACCTAGAACAGAGCTGCCCTTACTGTGC	13532
QY	2089	TGTTTCCCTAAGCAATTTATTAAGACCTACCTGGGTGCTGGGTTCACTGTGCTCTAGAAA	2148
Db	13533	TGTTTCCCTAAGCAATTTATTAAGACCTACCTGGGTGCTGGGTTCACTGTGCTCTAGAAA	13592
QY	2149	CCAAGAGGGTCCCGAGTCTTGACCTCTGACCCGCGCTGCGCCCAACAACCTTCTGACA	2208
Db	13593	CCAAGAGGGTCCCGAGTCTTGACCTCTGACCCGCGCTGCGCCCAACAACCTTCTGACA	13652
QY	2209	CACAGCGGTGGGGAGGCGGGAGAGACAGCTGGGACCCAGAACTGAGCCTGGAGGGATC	2268
Db	13653	CACAGCGGTGGGGAGGCGGGAGAGACAGCTGGGACCCAGAACTGAGCCTGGAGGGATC	13712
QY	2269	CGACAGAAAAGCTCAGGGCGGGTCTTCTCTCTGTGACCCGGGATTTGGAGCTATGCTGGATC	2328
Db	13713	CGACAGAAAAGCTCAGGGCGGGTCTTCTCTCTGTGACCCGGGATTTGGAGCTATGCTGGATC	13772
QY	2329	CACCATGTACTCAGGCAATGGTGGTTTTGAAACCCATPAACAAAGGCCCTTGTCAATCAGC	2388
Db	13773	CACCATGTACTCAGGCAATGGTGGTTTTGAAACCCATPAACAAAGGCCCTTGTCAATCAGC	13832
QY	2389	TCTTAAACAAGATATTTTGTATTTTATCTCTAACAACATATGAAGTTTATGGGCCCTTA	2448
Db	13893	TCTTAAACAAGATATTTTGTATTTTATCTCTAACAACATATGAAGTTTATGGGCCCTTA	13892
QY	2449	AGGAACCTTAGATCTTCTATTTGGGCTCTTCTGAGGTTTCAGAGAGGTAGTAACTTCC	2508
Db	13893	AGGAACCTTAGATCTTCTATTTGGGCTCTTCTGAGGTTTCAGAGAGGTAGTAACTTCC	13952
QY	2509	TTCAGGTCACACAGCAAGTCTGTGGGTGGCGAAGACGCTAGCGCTGGGCAATTCAATAC	2568
Db	13953	TTCAGGTCACACAGCAAGTCTGTGGGTGGCGAAGACGCTAGCGCTGGGCAATTCAATAC	14012
QY	2569	ATACACAGATGTGCTCCCTCTCTTGATGCTTGGACCCCTGGGGGCTTCAAGGCTTTGGGAC	2628
Db	14013	ATACACAGATGTGCTCCCTCTCTTGATGCTTGGACCCCTGGGGGCTTCAAGGCTTTGGGAC	14072
QY	2629	ATCTTTGTCTCAACCTCTCTCCCTAGATCACTGTGTGAAGGGTCCCTGTAGATATTTGTATC	2688
Db	14073	ATCTTTGTCTCAACCTCTCTCCCTAGATCACTGTGTGAAGGGTCCCTGTAGATATTTGTATC	14132
QY	2689	ACCATGCCCAATGTATATTAACAAGTACACACAGATGTACACACAGATGTACATATGCTCCAG	2748
Db	14133	ACCATGCCCAATGTATATTAACAAGTACACACAGATGTACACACAGATGTACATATGCTCCAG	14192
QY	2749	CCCCAGCTTGCAATACCTGCAACCTGCAACCCCAAGCCTTGGACCCCTGCTGCTGTGTGCTC	2808
Db	14193	CCCCAGCTTGCAATACCTGCAACCTGCAACCCCAAGCCTTGGACCCCTGCTGCTGTGTGCTC	14252
QY	2809	AAAGCACAAGCTCCAAACCTTGCTCTGTCTCCCTTCCCAACCAATGCTCTGAGCCTTCTGA	2868
Db	14253	AAAGCACAAGCTCCAAACCTTGCTCTGTCTCCCTTCCCAACCAATGCTCTGAGCCTTCTGA	14312
QY	2869	GCAAGACCAAGTACCTTGGCTGACACCGGTGTGTGGCCGCTTCAACCAAGGACACACCCCG	2928
Db	14313	GCAAGACCAAGTACCTTGGCTGACACCGGTGTGTGGCCGCTTCAACCAAGGACACACCCCG	14372
QY	2929	CCACCATGATCTTCGCTGTACACTATCATATAAAGTGGGTTTGTACAAA 2978	

```

DB          14373  CCACMANGATCTCCGTGTACACTATCAATAAAGTGGGTTTGTACAA 14422

RESULT 7
US-10-321-039-13
; Sequence 13, Application US/10321039
; Publication No. US20040014067A1
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Lukowiak, Andrew
; APPLICANT: Darvas, Nancy
; APPLICANT: Kurensky, David
; TITLE OF INVENTION: Amplification Methods and Compositions
; FILE REFERENCE: FORS-06960
; CURRENT APPLICATION NUMBER: US/10/321,039
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: 09/998,157
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/329,113
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/360,489
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 759
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (561)..(561)
; OTHER INFORMATION: n can be g or t.
US-10-321-039-13

```

	Only Match	33.0%	Score 1018.4	DB 16	Length 1021
	Best Local Similarity	99.8%	Pred. No. 8.6e-182		
	Matches 1019	Conservative	0	Mismatches 2	Indels 0
				Gaps	0
QY	665	CCACACCCGAGGCCAGAGCTGTCTGTGTGGGCGCAGAACTCCGCTACTTTCAGAGGT	724		
Db	1	CCACACCCGAGGCCAGAGCTGTCTGTGTGGGCGCAGAACTGCGCTTACTTGAGGT	60		
QY	725	CGGCGAAGAAAGAACCAACGTGACGAGATGTTCTAGTGCTTTTCAGATGGCCAA	784		
Db	61	CGGCGAAGAAAGAACCAACGTGACGAGATGTTCTAGTGCTTTTCAGATGGCCAA	120		
QY	785	TGCGACACGAGATGAGGCCCGCGCTGTGATCGCAAGATCTCCGTGCAATACGGTGA	844		
Db	121	TGCGACACGAGATGAGGCCCGCGCTGTGATCGCAAGATCTCCGTGCAATACGGTGA	180		
QY	845	TCCACCCCGAGGCTTTGTGATGCGCGCGTCAAGAGATGGAAGCCTATGGATGGTCT	904		
Db	181	TCCACCCCGAGGCTTTGTGATGCGCGCGTCAAGAGATGGAAGCCTATGGATGGTCT	240		
QY	905	CGGCTTTGCGCCCGCGCGCCCAAGCGTCAACATGACCTTCAAGTACATCAAGGCCA	964		
Db	241	CGGCTTTGCGCCCGCGCGCCCAAGCGTCAACATGACCTTCAAGTACATCAAGGCCA	300		
QY	965	TTTCGGAAAGGCGAGGCCCGCTGTGAGAGGCAAGATGCACATTCAGGTGACGAGGAT	1024		
Db	301	TTTCGGAAAGGCGAGGCCCGCTGTGAGAGGCAAGATGCACATTCAGGTGACGAGGAT	360		
QY	1025	GGGCGGGGCTTGGCGAGTGCCTTACAGGAGGTGGGCCCAATGCGCATGTGGGCAATTC	108		
Db	361	GGGCGGGGCTTGGCGAGTGCCTTACAGGAGGTGGGCCCAATGCGCATGTGGGCAATTC	420		
QY	1085	CCGACCGAGGCCCGCGCGAGAGTCTTTTTCACACACTTATGGACACGATGAGAGCC	1144		
Db	421	CCGACCGAGGCCCGCGCGAGAGTCTTTTTCACACACTTATGGACACGATGAGAGCC	480		
QY	1145	CGGCGGCTTGAGCTCCGCAATTCGTCTGCTTCTCACAGCTTTTCTTGAGTCCGCTTGTCC	120		
Db	481	CGGCGGCTTGAGCTCCGCAATTCGTCTGCTTCTCACAGCTTTTCTTGAGTCCGCTTGTCC	540		





```
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 25124
;; LENGTH: 753
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-25124
```

```
Query Match      23.8%; Score 734; DB 13; Length 753;
Best Local Similarity 99.1%; Pred. No. 2.3e-128;
Matches 745; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 695 GCGAGAGAACTCCGCCCTACTTCGAGGTGTGCGCCCAAGAAACCAACGTCGACGAGA 754
Db 753 GCGAGAGAACTCGCCCTTACTT-GAGGTGTGCGCCCAAGAAACCAACGTCGACGAGA 695
QY 755 TGTTCAGAGTGTCTTTCAGCATGGCCCAAGCTGCAACGAGATGAGCCCGCCCTGCATC 814
Db 694 TGTTCAGAGTGTCTTTCAGCATGGCCCAAGCTGCAACGAGATGAGCCCGCCCTGCATC 635
QY 815 GCAAGATCTCCGTGAGTACGGTGAAGCGCTTTCACCCCAAGCGCTTTCGATGCGCCG 874
Db 634 GCAAGATCTCCGTGAGTACGGTGAAGCGCTTTCACCCCAAGCGCTTTCGATGCGCCG 575
QY 875 TCAAGAGATGAGCGCCCTATGCGATGTGCGCCCTTGGCCCGCCCGCCGACGACCA 934
Db 574 TCAAGAGATGAGCGCCCTATGCGATGTGCGCCCTTGGCCCGCCCGCCGACGACCA 515
QY 935 GTGACCTCAAGTACATCAAGGCAAGGTCTTTCGGAAGGCGCAAGGCGCGTGAAGAGCA 994
Db 514 GTGACCTCAAGTACATCAAGGCAAGGTCTTTCGGAAGGCGCAAGGCGCGTGAAGAGCA 455
QY 995 AGTGCACCATTCAGTGAAGGAGGATGCTGCGGCGCGGCTTGGCCAGTGCCTTCAAGGAG 1054
Db 454 AGTGCACCATTCAGTGAAGGAGGATGCTGCGGCGCGGCTTGGCCAGTGCCTTCAAGGAG 395
QY 1055 GTGCGCCCAAGATGCGCATGTGCGCATCTCCCAACGAGCGCCCGGCGAGAGTCTGTTC 1114
Db 394 GTGCGCCCAAGATGCGCATGTGCGCATCTCCCAACGAGCGCCCGGCGAGAGTCTGTTC 335
QY 1115 ACAGACTTTAGGCAACAGACTGAGAGGCCCGCGGCGCTGCGCATTCGTCGTGC 1174
Db 334 ACAGACTTTAGGCAACAGACTGAGAGGCCCGCGGCGCTGCGCATTCGTCGTGC 275
QY 1175 TTCTCAGAGCTTTCCTGAGTCCGCTTGTTCACAGCTCTTGTGTTTCATCTCTGT 1234
Db 274 TTCTCAGAGCTTTCCTGAGTCCGCTTGTTCACAGCTCTTGTGTTTCATCTCTGT 215
QY 1235 GGGAGAGACATCTTCGAGCGCTCAAGAGTTTGGAGAGAGCTCAAGTTTACACTTCTCT 1294
Db 214 GGGAGAGACATCTTCGAGCGCTCAAGAGTTTGGAGAGAGCTCAAGTTTACACTTCTCT 155
QY 1295 CTTGGGCTTTGGAAGAAATGTTGATGCCAGAGGGGTGAAGATTCTGCGTCATATGAGGC 1354
Db 154 CTTGGGCTTTGGAAGAAATGTTGATGCCAGAGGGGTGAAGATTCTGCGTCATATGAGGC 95
QY 1355 TCTTCGAGACAGCTTCAGATGAAAAAGACACAGAAAGCCAGATGAGAAAGGTCTCTCT 1414
Db 94 TCTTCGAGACAGCTTCAGATGAAAAAGACACAGAAAGCCAGATGAGAAAGGTCTCTCT 35
QY 1415 CTCCTGGCATACACCCAGCTTGTGTTGGGCTG 1446
Db 34 CTCCTGGCATACACCCAGCTTGTGTTGGGCTG 3
```

```
US-10-027-632-25125/c
; Sequence 25125, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 25125
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-25125
```

```
Query Match      23.8%; Score 734; DB 13; Length 753;
Best Local Similarity 99.1%; Pred. No. 2.3e-128;
Matches 745; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 695 GCGAGAGAACTCCGCCCTACTTCGAGGTGTGCGCCCAAGAAACCAACGTCGACGAGA 754
Db 753 GCGAGAGAACTCGCCCTTACTT-GAGGTGTGCGCCCAAGAAACCAACGTCGACGAGA 695
QY 755 TGTTCAGAGTGTCTTTCAGCATGGCCCAAGCTGCAACGAGATGAGCCCGCCCTGCATC 814
Db 694 TGTTCAGAGTGTCTTTCAGCATGGCCCAAGCTGCAACGAGATGAGCCCGCCCTGCATC 635
QY 815 GCAAGATCTCCGTGAGTACGGTGAAGCGCTTTCACCCCAAGCGCTTTCGATGCGCCG 874
Db 634 GCAAGATCTCCGTGAGTACGGTGAAGCGCTTTCACCCCAAGCGCTTTCGATGCGCCG 575
QY 875 TCAAGAGATGAGCGCCCTATGCGATGTGCGCCCTTGGCCCGCCCGCCGACGACCA 934
Db 574 TCAAGAGATGAGCGCCCTATGCGATGTGCGCCCTTGGCCCGCCCGCCGACGACCA 515
QY 935 GTGACCTCAAGTACATCAAGGCAAGGTCTTTCGGAAGGCGCAAGGCGCGTGAAGAGCA 994
Db 514 GTGACCTCAAGTACATCAAGGCAAGGTCTTTCGGAAGGCGCAAGGCGCGTGAAGAGCA 455
QY 995 AGTGCACCATTCAGTGAAGGAGGATGCTGCGGCGCGGCTTGGCCAGTGCCTTCAAGGAG 1054
Db 454 AGTGCACCATTCAGTGAAGGAGGATGCTGCGGCGCGGCTTGGCCAGTGCCTTCAAGGAG 395
QY 1055 GTGCGCCCAAGATGCGCATGTGCGCATCTCCCAACGAGCGCCCGGCGAGAGTCTGTTC 1114
Db 394 GTGCGCCCAAGATGCGCATGTGCGCATCTCCCAACGAGCGCCCGGCGAGAGTCTGTTC 335
QY 1115 ACAGACTTTAGGCAACAGACTGAGAGGCCCGCGGCGCTGCGCATTCGTCGTGC 1174
Db 334 ACAGACTTTAGGCAACAGACTGAGAGGCCCGCGGCGCTGCGCATTCGTCGTGC 275
QY 1175 TTCTCAGAGCTTTCCTGAGTCCGCTTGTTCACAGCTCTTGTGTTTCATCTCTGT 1234
Db 274 TTCTCAGAGCTTTCCTGAGTCCGCTTGTTCACAGCTCTTGTGTTTCATCTCTGT 215
QY 1235 GGGAGAGACATCTTCGAGCGCTCAAGAGTTTGGAGAGAGCTCAAGTTTACACTTCTCT 1294
```

Db 214 GGGAGGACATCTCTGAGCTCAAGATTAGGAGAGACTCAAGTTACACCTTCCCTCT 155  
Qy 1295 CCTGGGTTTGAAGAATGTTGATGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACC 1354  
Db 154 CCTGGGTTTGAAGAATGTTGATGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACC 95  
Qy 1355 TCTGGGCAAGCCCTCAGAGTGAAGAAAGACAAGAGGCCAGATGAGAAAGTCTCTCT 1414  
Db 94 TCTGGGCAAGCTCAGAGTGAAGAAAGACAAGAGGCCAGATGAGAAAGTCTCTCT 35  
Qy 1415 CTCTGGCATTAACACCCAGCTTGTGTTGGGTG 1446  
Db 34 CTCTGGCATTAACACCCAGCTTGTGTTGGGGG 3

RESULT 11  
US-10-027-632-25123/c  
; Sequence 25123, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 25123  
; LENGTH: 753  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-25123

Query Match 23.8%; Score 734; DB 16; Length 753;  
Best Local Similarity 99.1%; Pred. No. 2.3e-128;  
Matches 745; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 695 GCGACGAGAACTCCGCTACTTCCAGGAGTGTGGCCCAAGAAACAACAAGTGAAGAGA 754  
Db 753 GCGACGAGAACTCCGCTACTT-GAGGTGTGGCCCAAGAAACAACAAGTGAAGAGA 695  
Qy 755 TGTCTAGGTGCTCTTTCAGCATGGCCCAAGCTCCACAGAGTGAAGCCCGCTTCATC 814  
Db 694 TGTCTAGGTGCTCTTTCAGCATGGCCCAAGCTCCACAGAGTGAAGCCCGCTTCATC 635  
Qy 815 GCAAGATCTCCGTGAGTGAAGGCTTTCACACCCCAAGGCTTTCGATGCGCCGG 874  
Db 634 GCAAGATCTCCGTGAGTGAAGGCTTTCACACCCCAAGGCTTTCGATGCGCCGG 575  
Qy 875 TCAAGAGATGAGCCCTATGAGTGTCTGCGCCCTTGCAGCCGCGCCCAAGCTCAACA 934  
Db 574 TCAAGAGATGAGCCCTATGAGTGTCTGCGCCCTTGCAGCCGCGCCCAAGCTCAACA 515  
Qy 935 GTGACCTCAAGTATCAAGGCAAGGCTCTTTCGGAAGGCCAGCCGCTGAGAGGACA 994  
Db 514 GTGACCTCAAGTATCAAGGCAAGGCTCTTTCGGAAGGCCAGCCGCTGAGAGGACA 455

Qy 995 AGTGACCATTCAGTGAAGCGAGGATGCTGGGGCGGGCTTGGCCAGTCCCTTCAGGAG 1054  
Db 454 AGTGACCATTCAGTGAAGCGAGGATGCTGGGGCGGGCTTGGCCAGTCCCTTCAGGAG 395  
Qy 1055 GTGGCCCAAGATGCCCATGTGGCATCTCCCAACCAAGGCCCGGAGAGCTTGTTC 1114  
Db 394 GTGGCCCAAGATGCCCATGTGGCATCTCCCAACCAAGGCCCGGAGAGCTTGTTC 335  
Qy 1115 ACAGACTTGAAGACACAGTGAAGGCCCCCGGCGCTGGCTCCGCAATTCGTCGCC 1174  
Db 334 ACAGACTTGAAGACACAGTGAAGGCCCCCGGCGCTGGCTCCGCAATTCGTCGCC 275  
Qy 1175 TTCTCAGACTTTCCTGAGTCCGCTGTCACAGCTCTTGTGATTCATCTCTCTGT 1234  
Db 274 TTCTCAGACTTTCCTGAGTCCGCTGTCACAGCTCTTGTGATTCATCTCTCTGT 215  
Qy 1235 GGGAGGACATCTCTGAGCCTCAAGATTAGGACAGACTCAAGTTACCTTCTCT 1294  
Db 214 GGGAGGACATCTCTGAGCCTCAAGATTAGGACAGACTCAAGTTACCTTCTCT 155  
Qy 1295 CCTGGGTTTGAAGAATGTTGATGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACC 1354  
Db 154 CCTGGGTTTGAAGAATGTTGATGCCAGAGGGGTGAGGATTTGCTGCTCATATGAGACC 95  
Qy 1355 TCTGGGCAAGCCCTCAGAGTGAAGAAAGACAAGAGGCCAGATGAGAAAGTCTCTCT 1414  
Db 94 TCTGGGCAAGCTCAGAGTGAAGAAAGACAAGAGGCCAGATGAGAAAGTCTCTCT 35  
Qy 1415 CTCTGGCATTAACACCCAGCTTGTGTTGGGTG 1446  
Db 34 CTCTGGCATTAACACCCAGCTTGTGTTGGGGG 3

RESULT 12  
US-10-027-632-25124/c  
; Sequence 25124, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 25124  
; LENGTH: 753  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-25124

Query Match 23.8%; Score 734; DB 16; Length 753;  
Best Local Similarity 99.1%; Pred. No. 2.3e-128;  
Matches 745; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 695 GCGACGAGAACTCCGCTACTTCCAGGAGTGTGGCCCAAGAAACAACAAGTGAAGAGA 754  
Db 753 GCGACGAGAACTCCGCTACTT-GAGGTGTGGCCCAAGAAACAACAAGTGAAGAGA 695

```
QY 755 TGTTCACGCGCTCTTACGAGTGGCCAAAGCTGCACACGAGATGAGCCCGCCCTGCATC 814
DB 694 TGTTCACGCGCTCTTACGAGTGGCCAAAGCTGCACACGAGATGAGCCCGCCCTGCATC 635
QY 815 GCAAGATCTCCGTGCAATGAGTGAAGCGCTTCCACCCCAAGCCCTTCTGATCGCCGCG 874
DB 634 GCAAGATCTCCGTGCAATGAGTGAAGCGCTTCCACCCCAAGCCCTTCTGATCGCCGCG 575
QY 875 TCAGAGATGAGCGCCCTTATGATGATGCTCTGCGCCCTTCCCGCCCGCCCAAGCTCAAC 934
DB 574 TCAGAGATGAGCGCCCTTATGATGATGATGCTCTGCGCCCTTCCCGCCCGCCCAAGCTCAAC 515
QY 935 GTGACCTCAAGTACATCAAGGCGCAAGGCTCTTGGGGAAGGCGAGCCCGTGAAGAGGACA 994
DB 514 GTGACCTCAAGTACATCAAGGCGCAAGGCTCTTGGGGAAGGCGAGCCCGTGAAGAGGACA 455
QY 995 AGTGACCATCCAGTGAAGGAGGATGCTGGGCGGGGCTTGGCCAGTGCCTTCAAGGAG 1054
DB 454 AGTGACCATCCAGTGAAGGAGGATGCTGGGCGGGGCTTGGCCAGTGCCTTCAAGGAG 395
QY 1055 GTGGCCCCAGATGCCCATCTGTGGCATCTTCCCAAGAGCCCGGAGAGATCTTGTTC 1114
DB 394 GTGGCCCCAGATGCCCATCTGTGGCATCTTCCCAAGAGCCCGGAGAGATCTTGTTC 335
QY 1115 ACAGACCTTAGGACCAAGACCTGGAGGCGCCCGGGGCTGGCCCTCCGACATTCGTCTGC 1174
DB 334 ACAGACCTTAGGACCAAGACCTGGAGGCGCCCGGGGCTGGCCCTCCGACATTCGTCTGC 275
QY 1175 TTCTCAGACCTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTTATCTCTCTGT 1234
DB 274 TTCTCAGACCTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTTATCTCTCTGT 215
QY 1235 GGAAGACATCTCTGCAAGCTTCAAGATGAGGAGACATCTCAATTCACCTTCTCT 1294
DB 214 GGAAGACATCTCTGCAAGCTTCAAGATGAGGAGACATCTCAATTCACCTTCTCT 155
QY 1295 CCTGGGGTTGGAAGAAATGTTGTCAGAGGGGTGAGGATTCGTCGATATGAGGCG 1354
DB 154 CCTGGGGTTGGAAGAAATGTTGTCAGAGGGGTGAGGATTCGTCGATATGAGGCG 95
QY 1355 TCTTGGGACCAAGCTTCAGATGAAAAGACACAGAAAGCCAGATGAGAAAGTCTCTCT 1414
DB 94 TCTTGGGACCAAGCTTCAGATGAAAAGACACAGAAAGCCAGATGAGAAAGTCTCTCT 35
QY 1415 CTCCTGGCATTAACCCAGCTTGTGGGTG 1446
DB 34 CTCCTGGCATTAACCCAGCTTGTGGGTG 3

RESULT 13
US-10-027-632-25125/c
; Sequence 25125, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
```

```
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25125
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-25125

Query Match      23.8%; Score 734; DB 16; Length 753;
Best Local Similarity 99.1%; Pred. No. 2,3e-128;
Matches 745; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 695 GCGAGGAGAACTCCGCTTACTTCAAGGTGTGTGCGCCAAAGAAACCAAGTGAAGAGA 754
DB 753 GCGAGGAGAACTCCGCTTACTT-GAGTGTGCGCCAAAGAAACCAAGTGAAGAGA 695
QY 755 TGTTCACGCGCTCTTCAAGCATGAGCCCAAGCTGCACAGAGATGAGCCCGCCCTGCATC 814
DB 694 TGTTCACGCGCTCTTCAAGCATGAGCCCAAGCTGCACAGAGATGAGCCCGCCCTGCATC 635
QY 815 GCAAGATCTCCGTGCAATGAGTGAAGCGCTTCCACCCCAAGCCCTTCTGATCGCCGCG 874
DB 634 GCAAGATCTCCGTGCAATGAGTGAAGCGCTTCCACCCCAAGCCCTTCTGATCGCCGCG 575
QY 875 TCAGAGATGAGCGCCCTTATGATGATGCTCTGCGCCCTTCCCGCCCGCCCAAGCTCAAC 934
DB 574 TCAGAGATGAGCGCCCTTATGATGATGATGCTCTGCGCCCTTCCCGCCCGCCCAAGCTCAAC 515
QY 935 GTGACCTCAAGTACATCAAGGCGCAAGGCTCTTGGGGAAGGCGAGCCCGTGAAGAGGACA 994
DB 514 GTGACCTCAAGTACATCAAGGCGCAAGGCTCTTGGGGAAGGCGAGCCCGTGAAGAGGACA 455
QY 995 AGTGACCATCCAGTGAAGGAGGATGCTGGGCGGGGCTTGGCCAGTGCCTTCAAGGAG 1054
DB 454 AGTGACCATCCAGTGAAGGAGGATGCTGGGCGGGGCTTGGCCAGTGCCTTCAAGGAG 395
QY 1055 GTGGCCCCAGATGCCCATCTGTGGCATCTTCCCAAGAGCCCGGAGAGATCTTGTTC 1114
DB 394 GTGGCCCCAGATGCCCATCTGTGGCATCTTCCCAAGAGCCCGGAGAGATCTTGTTC 335
QY 1115 ACAGACCTTAGGACCAAGACCTGGAGGCGCCCGGGGCTGGCCCTCCGACATTCGTCTGC 1174
DB 334 ACAGACCTTAGGACCAAGACCTGGAGGCGCCCGGGGCTGGCCCTCCGACATTCGTCTGC 275
QY 1175 TTCTCAGACCTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTTATCTCTCTGT 1234
DB 274 TTCTCAGACCTTCTGAGTCCGCTTGTCCACAGCTCTTGTGTGTTTATCTCTCTGT 215
QY 1235 GGAAGACATCTCTGCAAGCTTCAAGATGAGGAGACATCTCAATTCACCTTCTCT 1294
DB 214 GGAAGACATCTCTGCAAGCTTCAAGATGAGGAGACATCTCAATTCACCTTCTCT 155
QY 1295 CCTGGGGTTGGAAGAAATGTTGTCAGAGGGGTGAGGATTCGTCGATATGAGGCG 1354
DB 154 CCTGGGGTTGGAAGAAATGTTGTCAGAGGGGTGAGGATTCGTCGATATGAGGCG 95
QY 1355 TCTTGGGACCAAGCTTCAGATGAAAAGACACAGAAAGCCAGATGAGAAAGTCTCTCT 1414
DB 94 TCTTGGGACCAAGCTTCAGATGAAAAGACACAGAAAGCCAGATGAGAAAGTCTCTCT 35
QY 1415 CTCCTGGCATTAACCCAGCTTGTGGGTG 1446
DB 34 CTCCTGGCATTAACCCAGCTTGTGGGTG 3

RESULT 14
US-10-029-386-20951
; Sequence 20951, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
```

APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: AEMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Autonom Sequence Listing Engine vers. 1.1  
SEQ ID NO 20951  
LENGTH: 696  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL022334.1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.6  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
OTHER INFORMATION: EST HUMAN HIT: B1754083.1, EVALUATE 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: P10301, EVALUATE 4.00e-10  
OTHER INFORMATION: NT HIT: g16168170, EVALUATE 0.00e+00  
US-10-029-386-20951

Query Match 22.5%; Score 694.4; DB 15; Length 696;  
Best Local Similarity 99.9%; Pred. No. 6.2e-121;  
Matches 695; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 482 GGGATGCTTCATCTGCTGTTCAAGCTGATTAACCGGAGTCTTCGATGAGTCAAC 541  
Db 1 GGGATGCTTCATCTGCTGTTCAAGCTGATTAACCGGAGTCTTCGATGAGTCAAC 60  
Qy 542 GCTTCAGAGCAGATCTTGAAGTCAAGTCTGCTTGAAGAACAAAGCAAGAGGCGG 601  
Db 61 GCTTCAGAGCAGATCTTGAAGTCAAGTCTGCTTGAAGAACAAAGCAAGAGGCGG 120  
Qy 602 CGGAGTCCCATGATCTGCTGCAAGAAAGCAAGCAAGGAGTGGCGCCGAG 661  
Db 121 CGGAGTCCCATGATCTGCTGCAAGAAAGCAAGCAAGGAGTGGCGCCGAG 180  
Qy 662 TGCCACCAACGAGGCGGAGTGTGTGCGGCGAAGAACTCCGCTTACGAGG 721  
Db 181 TGCCACCAACGAGGCGGAGTGTGTGCGGCGAAGAACTCCGCTTACGAGG 240  
Qy 722 TGTGGGCAAGAAAGCAAGCTGAGAGATTTCTAGTGTCTTCAAGCATGGCCA 781  
Db 241 TGTGGGCAAGAAAGCAAGCTGAGAGATTTCTAGTGTCTTCAAGCATGGCCA 300  
Qy 782 AGCTGCCACAGAGATGAGCCCGGCTGATCGCAATCTCCGTCAGTACGCTGAG 841  
Db 301 AGCTGCCACAGAGATGAGCCCGGCTGATCGCAATCTCCGTCAGTACGCTGAG 360  
Qy 842 CTTTCACCAAGGCTTCTGATGCGCGGTCAGAGAGATGAGAGCTTATGAGCATG 901  
Db 361 CTTTCACCAAGGCTTCTGATGCGCGGTCAGAGAGATGAGAGCTTATGAGCATG 420  
Qy 902 TCTGCGCTTGGCGCGCGCGCGCGAGCTGACAGTGAAGTCAAGTCAAGGCGAAG 961  
Db 421 TCTGCGCTTGGCGCGCGCGCGCGAGCTGACAGTGAAGTCAAGTCAAGGCGAAG 480  
Qy 962 TCTGCGCGAGGCGAGGCGCGGTCAGAGAGGAGCAATGCAATCCAGTGAAGGAGG 1021  
Db 481 TCTGCGCGAGGCGAGGCGCGGTCAGAGAGGAGCAATGCAATCCAGTGAAGGAGG 540  
Qy 1022 CTGGGCGGAGGCTTGGCGAGTGTCTTCAAGAGAGTGGCCAGATGTCGAGAT 1081  
Db 541 CTGGGCGGAGGCTTGGCGAGTGTCTTCAAGAGAGTGGCCAGATGTCGAGAT 600  
Qy 1082 CTCCCACTGAGAGGCGCGAGAGTCTTGTTCACAGACTTTAGGACCAAGACTGAGGC 1141  
Db 601 CTCCCACTGAGAGGCGCGAGAGTCTTGTTCACAGACTTTAGGACCAAGACTGAGGC 660  
Qy 1142 CCGCGGAGGCTGCGCTCGGACATTCGTCGCTTC 1177

Db 661 CCGCGGAGGCTGCGCTCGGACATTCGTCGCTTC 696

RESULT 15  
US-09-918-715-292  
Sequence 292, Application US/09918715  
Publication No. US20030017157A1  
GENERAL INFORMATION:  
APPLICANT: Brad St. Croix  
APPLICANT: Kenneth Kinzler  
TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
FILE REFERENCE: 1107.00134  
CURRENT APPLICATION NUMBER: US/09/918,715  
CURRENT FILING DATE: 2001-08-01  
PRIOR APPLICATION NUMBER: 60/222,599  
PRIOR FILING DATE: 2000-08-02  
PRIOR APPLICATION NUMBER: 60/224,360  
PRIOR FILING DATE: 2000-08-11  
PRIOR APPLICATION NUMBER: 60/282,850  
PRIOR FILING DATE: 2000-04-11  
NUMBER OF SEQ ID NOS: 358  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 292  
LENGTH: 3020  
TYPE: DNA  
ORGANISM: Mouse  
US-09-918-715-292

Query Match 22.5%; Score 693.2; DB 13; Length 3020;  
Best Local Similarity 87.0%; Pred. No. 1.8e-120;  
Matches 774; Conservative 0; Mismatches 113; Indels 3; Gaps 1;

Qy 200 AGCCCGAGCCATGATGAGACTTTGTCAAGCGGAACTGACGCTCAGTGTGCCGCA 259  
Db 365 AGCCCGAGCCATGATGAGACTTTGTCAAGCGGAACTGACGCTCAGTGTGCCGCA 424  
Qy 260 AAAAATCATACCGATGCTGTGCTGCTGCTTCCGGTGGCAAGTCTCATCGTGT 319  
Db 425 AAAAATCATACCGATGCTGTGCTGCTGCTTCCGGTGGCAAGTCTCATCGTGT 484  
Qy 320 CTCGCTTCCTATGAGCGCGCTTGTGAGGACAGTACACACCACCATGAGGACTTCCACC 379  
Db 485 CTCGCTTCCTATGAGCGCGCTTGTGAGGACAGTACACACCACCATGAGGACTTCCACC 544  
Qy 380 GTAAAGTATCAACATCCGCGGCAATGTACAGCTGACATCTGATCTTGGCA 439  
Db 545 GTAAAGTATCAACATCCGCGGCAATGTACAGCTGACATCTGATCTTGGCA 604  
Qy 440 ACCACCTTTCGCGGCAATGCGAGCTGTCTCATCTTCAAGGAGATCTTCAATCTG 499  
Db 605 ACCACCTTTCGCGGCAATGCGAGCTGTCTCATCTTCAAGGAGATCTTCAATCTG 664  
Qy 500 TGTTCAGCTGATTAACCGGAGTCTTCATGAGAGTCAAGGCTTCAAGGAGATCC 559  
Db 665 TGTTCAGCTGATTAACCGGAGTCTTCATGAGAGTCAAGGCTTCAAGGAGATCC 724  
Qy 560 TGGAGTCAAGTCTTGCCTGAGAAACAAGCAAGAGAGGCGAGAGTGCCTGCAATG 619  
Db 725 TGGAGTCAAGTCTTGCCTGAGAAATAAACAAGAGAGGCGAGAGTGCCTGCAATG 784  
Qy 620 TCTGTGCAACAAGAGCAAGCAAGCGAGCTGTGCTGCGGAGTGTCCACACCGAGGCG 679  
Db 785 TCTGTGCAACAAGAGCAAGCAAGCGAGCTGTGCTGCGGAGTGTCCACACCGAGGCG 844  
Qy 680 AGCTGTGCTGTGCGGCGAGAGAACTCCGCTTCTTCAAGTGTGCGGCAAGAGACA 739  
Db 845 AGCTGTGCTGTGCGGCGAGAGAACTCCGCTTCTTCAAGTGTGCGGCAAGAGACA 904  
Qy 740 CCAAGTGAAGAGATGTTCTACGCTTCTTCAAGATGCGCAAGCTGCGACAGAGATGA 799  
Db 905 CCAAGTGAAGAGATGTTCTACGCTTCTTCAAGATGCGCAAGCTGCGACAGAGATGA 964



Db 181 CCTGGG-CTTCAGGGCTTTGGGACATCTGTCTCCTCAACCCCTCTCCCTAGATCAGTCTGT 239

QY 2664 GAGGGTCCCTGATGATTTGTGACACATGCCATGTATATACAAGTACACACAGATGT 2723  
DB 240 GAGGGTCCCTGATGATTTGTGACACATGCCATGTATATACAAGTACACACAGATGT 299  
QY 2724 ACACACAGATGTACACATGTCTCCAGCCAGCTCTGCATATCTGCACCTGCACCCAGCC 2783  
DB 300 ACACACAGATGTACACATGTCTCCAGCCAGCTCTGCATATCTGCACCTGCACCCAGCC 359  
QY 2784 TTGGCCCCCTGCTGTCTGTGTCTCAAGAGACAGCTCCACCTGCTCTGTCTCTTC 2843  
DB 360 TTGGCCCCCTGCTGTCTGTGTCTCAAGAGACAGCTCCACCTGCTCTGTCTCTTC 419  
QY 2844 CCCACCATGCTGCTGAGCTCTTCTGAGACAGACAGTACCTTGCTGACACCGGTGTGG 2902  
DB 420 CCCACCATGCTGCTGAGCTCTTCTGAGACAGACAGTACCTTGANTAAACGGGTGTGG 478

RESULT 2  
US-09-053-374A-1  
; Sequence 1, Application US/09053374A  
; Patent No. 6462177  
; GENERAL INFORMATION:  
; APPLICANT: YEN, KWANG-MU  
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC.  
; STREET: ONE AMGEN CENTER DRIVE  
; CITY: THOUSAND OAKS  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/053,374A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOK, ROBERT R.  
; REGISTRATION NUMBER: 31,602  
; REFERENCE/DOCKET NUMBER: A-514  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1841 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 255..1097  
; US-09-053-374A-1

Query Match 13.5%; Score 416; DB 4; Length 1841;  
Best Local Similarity 65.9%; Pred. No. 8,2e-73;  
Matches 668; Conservative 0; Mismatches 310; Indels 36; Gaps 3;

QY 25 AGCCGCGCCAGCCCGGCGTCCGAGACGCGAGGGAGATCCCGCGCAGTGAACCGG 84  
DB 84 AGCCGCGCTAGCCCGAGCCGAGCCGAGCGAAGCCGAGCCCGAGCCCGCGCC 143  
QY 85 GAGCCACACACACTCTGTGAGAGCTGTGGCGGCTGTGAGACAGCAGGACGCTCCCGCAGCTC 144  
DB 144 CAGCCCGAGACAGGCTCTCAACCGCTCAACCCCGGTGCAACCCAGGACCCCTCAGCG 203  
QY 145 CCGGCGCTTCCAGGAGCTCTGTGAGCGGTGCAAGAGCCCGGCGCCGATTCAGCGCC 204  
DB 204 CTCTCTGCCCTTCTCTCGGCGCCGCGCCGCTCTGCGGCGCCCTCTGCGCAATGAACTG 263

QY 205 CGAGCATATGAAACTTTTCTCAACCGGAACTGCAAGCTCAAGTGTGCGCCGCAAAAC 264  
DB 264 GCGGAGATGATMAAAGATGTGCCGAGCGACTGAGAGCTGATATCCCGCCAAAGAC 323  
QY 265 TCATACCGCATGT 324  
DB 324 TGTATCGCATGT 383  
QY 325 TTCTCAATGCGCGCTTTGAGAGCAGTACACACCCATGAGAGTCTTCCACCGTAAG 384  
DB 384 TTCTCAACCGCGCTTTGAGAGCAGTACACACCCATGAGAGTCTTCCACCGTAAG 443  
QY 385 GTATCAACATCCCGGCGGAGCATGTACAGCTTCAATCTGTGATATCTGTGCAACAC 444  
DB 444 TTCTCAATCCCGGCGGAGCATGTACAGCTTCAATCTGTGATATCTGTGCAACAC 503  
QY 445 CCTTCCCGCATGCGAGGCTGTCAATCTTCAACAGGAGATGTCTTATCTGCTGTTC 504  
DB 504 CGGTCCCGCATGCGGCGCTTCTCAATCTTCAACAGGAGATGTCTTATCTGCTGTTC 563  
QY 505 AGCTTGATTAACCGGAGTCTTGTGATGAGTCAAGCGCTTTCAGAAACAGATCTGTGAG 564  
DB 564 AGCTTGATTAACCGGAGTCTTGTGATGAGTCAAGCGCTTTCAGAAACAGATCTGTGAG 623  
QY 565 GTCAAGCTGTGTGAAGAAAGAAAGAAAGAGAGGCGGAGCTGCCATGTATCTGT 624  
DB 624 ACCAAGCTTGTGTGAAGAAAGAAAGAAAGAGAGGCGGAGCTGCCATGTATCTGT 683  
QY 625 GCGAACAAAGACGACAGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTG 684  
DB 684 GCGAACAAAGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTG 740  
QY 685 CTGTGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTG 741  
DB 741 CTGTGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTG 800  
QY 742 AACGGAACGAGATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 801  
DB 801 AACGGAACGAGATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 860  
QY 802 CCGGCGCTGATCGCAAGATCTCGGTGAGTACGAGTACGAGTGTGAGTGTGAGTGTGAGTGTGAG 857  
DB 861 CCGGCGCTGATCGCAAGATCTCGGTGAGTACGAGTACGAGTGTGAGTGTGAGTGTGAGTGTGAG 920  
QY 858 -----CTTGTGATGCGCGCGCTGTCAAGAGATGAGCGC 891  
DB 921 CGAAGCAAGAGCTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGG 980  
QY 892 TATGATGATGTGTGCGCTTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTG 951  
DB 981 TTTGGCATGTGTGAGTGTGCGCTTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGG 1040  
QY 952 AAGGCAAGTGTCTTGTGGAAGGCTGAGGCGGCTGTGAGAGGAGCAAGTGCACCATC 1005  
DB 1041 CGCAGAAAGGCTGAGGCGGCTGTGAGAGGAGCAAGTGCACCATC 1094

RESULT 3  
US-09-053-374A-4  
; Sequence 4, Application US/09053374A  
; Patent No. 6462177  
; GENERAL INFORMATION:  
; APPLICANT: YEN, KWANG-MU  
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC.  
; STREET: ONE AMGEN CENTER DRIVE  
; CITY: THOUSAND OAKS  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,374A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-514  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1689 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 132..971  
US-09-053-374A-4

Query Match 11.2%; Score 346.6; DB 4; Length 1689;

Best Local Similarity 65.0%; Pred. No. 3.5e-59;  
Matches 558; Conservative 0; Mismatches 274; Indels 27; Gaps 2;

174 TGCAGAGGCGCCGCGCCGATCCAGCCCGGAGCCATGATGAAGACTTGTCCAGCG 233  
110 TGCCTGGGCGCCCTCTGCCAATGAACCTGGCCGATGATCAAGAGATGTGCCAAG 169  
234 GAACCTGACGCTGACGTGTGCGCCGCAAAACTCATACCGATGTGTGTGCGCTC 293  
170 CGACTCTGAAGTGAATGATCCGCGCAAGACTGCTACAGATGTGATCTCGGCTATC 229  
294 TCGGCTGGGCAAGACTCATGCTGTCTGCTCTCTCTCAATGGCGCTTTGAGACAGTA 353  
230 CAAGTGGGCAAGACGGCGATCGTGTGCGCTCTCTCAAGGCGCGCTTTCAGAGCGCTTA 289  
354 CAGACCCACCATGAGGAGCTTCCACCGTAAGTATACAACTCGCGCGGCGAATGTACCA 413  
290 CACCCCTACATTAAGACTTCCACCGAAAGTTTACTGATCCGCGCGAAGTCTACCA 349  
414 GCTGACATCTGTGATACCTGTGCAACACCCCTTCCCGCATGCGCAGGCTGTCCAT 473  
350 GTTGAACATATGAGACATCTGGCAATCATCCGTTTCCGCGCATGGCGGCTCTCAT 409  
474 CTTCAAGGGGATGTTCTTCAATCTGTGTGTTACCTGATTAACCGGAGTCTTTCATGA 533  
410 CCTCACAGGAGACGTTTTCATTTGTGTGTTTCAAGCTTAACACCGCATCTCTTCAGAGA 469  
534 GGTCAAGGCGCTTTCAGAGAGATCTTGAGGTCAAGCTTGCCTGAGAAACAAGACCA 593  
470 GGTGCAAGGCTTCAACAGCAGATCTTGAACACCAAGTCTGTCTCAAGAAACAACCA 529  
594 GAGAGCGCGGAGCTGCGCATGTGATCTGTGCAACAGAAACAACAACGCGAGCTGTG 653  
530 AGAAGATGTGACGTCGCGCTGTGATTTGCGGTAAACAAGGAGGACCGGAGCTTCTACG 589  
654 CCGGCAAGGTGCGCACACCGAGGCGGAGCTGTGTGTCGGGCGACGAGAACTCCGCTTA 713  
590 CGAAGTGAAGAGCGGAGATTTGACACAGCTGTGTGGGCATGACCTTCAGCGTTTGCTTA 649  
714 CTTGAGAGTGTGCGCAAGAGAAACAACAAGTGAAGATGTTCTTACGTGTCTTTCAG 773  
650 CTTGAGATCTTGGCGCAAGAGAAAGAGAGAGCTGAGCAAGATGTTCCGTGGCTCTTTCG 709  
774 CATGGCAAGCTGCGACAGAGATGAGCCCGCTGATGCAAGATCTTCGTCAGATA 833  
710 CATGGCAAGCTGCGACAGAGATGAGCCCGCTGATGCAAGATGTCGTGTGTCAGATA 769

834 CGGTAGCGCTTCCAC-----CCAGGCGCTTGTGATGCGCCCGCTCA 877  
770 CTGTGACGTGTGACCAAAAAAGCTGTGAGGAACAAGAGTTTCTGCTGGGCGACGG 829  
878 AGAGATG-----GAGCGCTATGCGATGTCGCTTCCGCGCGCGCCCGAG 926  
830 AGGTGGGGGACACCGAGAGATGCTTTGGCATCTTGGCGCGCTTGTGCTCGACAGCTAG 889  
927 CGTCAACAGTACCTCAAGTACATCAAGGCCAAGCTCTTGGGAAAGGCCCGCTGA 986  
890 CGTCAATAGGCACTCATGTACATTCGTGAGAAAAACAGTGTACAGCCAGGCTTAAGA 949  
987 GAGGACAAGTGACCATC 1005  
950 CAAGAGCGCTGTGTATC 968

#### RESULT 4

US-09-053-374A-3

Sequence 3, Application US/09053374A

Patent No. 6462177

GENERAL INFORMATION:

APPLICANT: YEN, KWANG-MU

TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSER: AMGEN INC.

STREET: ONE AMGEN CENTER DRIVE

CITY: THOUSAND OAKS

STATE: CA

COUNTRY: US

ZIP: 91320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/053,374A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: COOK, ROBERT R.

REGISTRATION NUMBER: 31,602

REFERENCE/DOCKET NUMBER: A-514

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3986 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-09-053-374A-3

Query Match 7.6%; Score 234.6; DB 4; Length 3986;  
Best Local Similarity 67.5%; Pred. No. 5e-37; Indels 37; Gaps 4;

Matches 407; Conservative 0; Mismatches 159; Indels 37; Gaps 4;

436 GGCACACACCCCTTCCCGCATGCGAGCTGTCAATCTTCACAGGGAGTCTTCAATC 495  
1226 GGCACACCTCACTTCTTCTTCTTCTGTGTCGCTGCGCCCTC-TAGAGAGATTTCATC 1284  
496 CTGTGTTCAGCTGTGATTAACCGGAGTCTTTCATGATGATCAAGCGCTTCAAGACAG 555  
1285 CTGTGTTCAGCTGTGATTAACCGGAGTCTTTCATGATGATGATCAAGCGCTTCAAGACAG 1344  
556 ATCCGAGGATCAAGTCTGCTGAGAAACAAGCAAGAGGCGGCGAGCTGCGCATG 615  
1345 ATCCGAGGATCAAGTCTGCTGAGAAACAAGCAAGAGGAGAGAGAGAGAGAGAGAGAG 1404  
616 GTCAATCTGTGCAACAAGAGCAAGCGAGAGTGTGCGCCAGAGTGTGCGCCAGACAG 675  
1405 GTCAATCTGTGCAACAAG 1461

OY	676	GGCGAGCTGCTGATGTCGGGCGA---CGAGAACTCCGCTACTTCGAGATGTCGGCCAAG	732
Db	1462	ATCGAGCAGCTGATGTCGGGCGACGACCCCGACGCTGCGCTACTTCGAGATCTCGGCAAG	1521
OY	723	AAGAACACCAACGTGAGACGATGTTCTTACGTGCTCTTCAGCATGGCCAAAGCTGCGCAAC	792
Db	1522	AAGAACAGAGCTGAGACCAAGATGTTCCGGCGGCTCTTCGCGATGGCCAAAGCTGCGCAGC	1581
OY	723	GAGATGAGCCCGCCCTCGATGCGAAGATCTCCGTGCAGTACGTGACGCTTCCACCCC	852
Db	1582	GAGATGAGCCCGACACTTGACCGCAAGCTTCGTGCAGTACTGCGAAGTGTCTGCAAG	1641
OY	853	AGGCC-----CTTCTGCATGCGCGGGTCAAGSAG	882
Db	1642	AAGGCGCTGCGGAAACAAGAGCTGCTGCGGGCCGGCAAGGGCGGGCGGCGGCAACCG	1701
OY	883	ATGACGCGCTATGCGATGATGTCGCGCTTCGCGCCGCGCCCGCAGCGTCAACAGTGAACCTC	942
Db	1702	GAGCAGCGCTTTGGCATGTCGTCACCTTCGCGGGCGGGCCGACAGGTACACAGGGAACCTC	1761
OY	943	AAGTTCATCAAGGCCAAGGCTCTTTCGGGAAGGCCAGGCCGCTGAGAGGGAACAAGTGCACC	1002
Db	1762	ATGTACATCCCGAGAAGGCCACGCGCCGCGACGCGCAAGAGCAACAAGAACGCTGCGTGC	1821
OY	1003	ATC 1005	
Db	1822	ATC 1824	

RESULT 5  
US-09-053-374A-6  
: Sequence 6, Application US/09053374A  
: Patent No. 6462177

```

1 GENERAL INFORMATION:
2 APPLICANT: YEN, KWANG-MU
3 TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
4 NUMBER OF SEQUENCES: 9
5 CORRESPONDENCE ADDRESS:
6 ADDRESSEE: AMGEN INC.
7 STREET: ONE AMGEN CENTER DRIVE
8 CITY: THOUSAND OAKS
9 STATE: CA
10 COUNTRY: US
11 ZIP: 91320
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: Patent In Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/09/053,374A
19 FILING DATE:
20 CLASSIFICATION: 435
21 ATTORNEY/AGENT INFORMATION:
22 NAME: COOK, ROBERT R.
23 REGISTRATION NUMBER: 31,602
24 REFERENCE/DOCKET NUMBER: A-514
25 INFORMATION FOR SEQ ID NO: 6:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 3079 base pairs
28 TYPE: nucleic acid
29 STRANDEDNESS: single
30 TOPOLOGY: linear
31 MOLECULE TYPE: DNA (genomic)
32 OS-09-053-374A-6

```

Query Match	6.5%	Score 199.6	DB 4	Length 3079
Best Local Similarity	56.0%	Pred. No. 3.5e-30		
Matches 558	Conservative	0	Mismatches 274	Indels 164
				Gaps 3

QY 174 TGCCAGAGGCCCGGCCCATTTCCAGCCCCGAGCCATGATGAGAGCTTTGTCCAGCGG 233

Db	281	TTCCCTGGGGCCCTCTGCGCCATGAACTGGCCGGAGATCAAGAAATGNGCCAA	344
QY	234	GAACCTGACGCTCAAGTGGCCCCGCAAAAATCTATACCGCATAGTGTGCTGGTGCCTC	293
Db	341	CGACTCTGAACTGAGTATCCCGCCAAAGACTGTCAAGAAATGTCATCTCGGCTCATTC	400
QY	294	TGCGGTGGGCAAGAGCTCATGTGTCTGCTTCTCAATGGCCGGCTTGAAGACAGTA	353
Db	401	CAAAATGGGCAAGAGGCGCATCTGTGTGGCTTCTCACGGGCCGCTTGAAGACGCTTA	460
QY	354	CACACCCACATCGAGGACTTCCACCGTAAGGTATCAACATCCGCGGACATGTACCA	413
Db	461	CACCCCTACATTTGAAGACTTCCACCGAAAGTTTACTGATCCGGGGGAAGTACCA	520
QY	414	GCTGCACATCCCTGATATCTCTGGCAACACCCCTTCCCGGCATAGCGAGGCTGCAT	473
Db	521	GTTCGACATATCTGACACATCTGGCAATCATCTGTTTCCCGCATAGCGGCGCTCTCAT	580
QY	474	CCTCAC-----	479
Db	581	CCTCACAGGTGAGTGGGGGACCGACAGGACCGTGGGAGGAAATCTGCGGGAGCGAT	640
QY	480	-----	479
Db	641	GGGGCGGTGTGTGCTTGGGGCTGTGCTGTCTGTGCTCGCTCCGCTTGGCAGTGCCCT	700
QY	480	-----AGGAGATGCTTCAATCTGTGTGTTCAGCTCGATAC	516
Db	701	CACCTTCACTGTTCCTTGTGAAGACGTTTTCATCTGTGTGTTCAGCTTAAACAAC	760
QY	517	CGGAGTCTTTCGATGAGGTCAAGGCGCTTCAAGACAGATCCGTGAAGGTCAAGTCTGC	576
Db	761	CGGACTCTTTCGAGGAGGTGAAAGGCTCAAAACGACGATCTTAGAACCAAGTCTGT	820
QY	577	CTGAAGAACAAAGACCAAGAGCGCGGAGCTGCCATGTTCATGTGTGGCAACAAGAC	636
Db	821	CTCAAGAACAAACCAAGAAAGATGTGACGTGCGCTGTCTATTTCGATTAACAAGGG	880
QY	637	GACCACGGCGAGCTGTGCGCGCAGGTGCCACACCGAGCGCGAGCTGTGTCTGGG	696
Db	881	GACCGGCGCTTACCGGAAATGTAGAGCGGGAGATTGACAGCATGTGTGGCGATAC	940
QY	697	GACGAAACTCCGCTACTTTCGAGGTGTCCGCCAAGAAACACCAACGTGACAGATG	756
Db	941	COTGAGCGTGTGCTACTTTCGAGATCTCGGCCAAGAAATAGAGCTGACCAAGATG	1000
QY	757	TTCTAAGTCTTTCAGCATGAGCCAAAGTGCACACGAGATAGCCCCGCTGTGATGC	816
Db	1001	TTCCGTGGCTCTTTCGCATGCGCAAGCTGTCTAGCGAGTAGAGCCTGACTTGCACGC	1060
QY	817	AAGATCTCCGTCAGTACGATGAGCGCTTCCAC-----CCAGGCCCTT	860
Db	1061	AAGGTGTCTGACGATCTGTGAGGTGTCTGACAAAAAGGCTCTGAGGAACAAGAGCTT	1120
QY	861	CTGCATGCGCGCGTCAAGAGATG-----GACGCTATAGCATGTCTCGGCC	909
Db	1121	CTGCGTGGGCAAGGAGGAGGTGGGGCGGACCAAGAGATGCTTTTGGATCTTGGCGCC	1180
QY	910	TTTCGCCGCGCCCGCAGGCTCAACATGTGACTTCAAGTACATCAAGGCCAGGTCTTGG	969
Db	1181	TTTCTGTCAACCTTAAGGTGACATAGGACCTTCAATGTACATTCGTGAGAAAAACAGTGT	1240
QY	970	GAAGCCAGGCGCGTGAAGAGGCAAGTGCACCATC 1005	
Db	1241	AGCAGCCAGGCTTAAGACAAAGAGCGCTGTCTATC 1276	

RESULT 6  
US-09-620-312D-945  
; Sequence 945, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom

APPLICANT: Tang, Y. Tom

```

Db      640 AACGTGAAGAGGCTTTTCAGAGGCTGCTCAACCTGGAGAAAGC 682

RESULT 7
US-09-621-976-17745
; Sequence 1.7745, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 17745
;
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17745

Query Match      3.9%; Score 119.4; DB 4; Length 432;
Best Local Similarity 59.1%; Pred. No. 1.1e-14;
Matches 204; Conservative 0; Mismatches 141; Indels 0; Gaps 0

Db      262 AACTCATATCCGCATGATGATGCTGGATGCTCCCTCGGAGTGGGCAAGAGCTTCATCGTGTCT 321
Oy      262 AACTCATATCCGCATGATGATGCTGGATGCTCCCTCGGAGTGGGCAAGAGCTTCATCGTGTCT 321
Db      12 AAACATTACCGGGTGGCGGTGTTGGGGCTGGCGGGTGTGGCAAGAGCTCCCTGGTGTGG 61
Oy      322 CGCTTCTCAATGCGCGCTTTGAGAGCAAGTACACACCACCATCGAGAGACTTCCACCGT 381
Db      62 AGGTTTGAAGGACCATTTCCGGGAGAGCTACATCCGACCGATGGAAACACTTACCGG 121
Oy      62 AGGTTTGAAGGACCATTTCCGGGAGAGCTACATCCGACCGATGGAAACACTTACCGG 121

```

```

OY      382  AAGGATATCAACATCCGGGGGCGACATGTAAACGACTCGACATCTGGATATCTTGGCAAC  441
Db      122  CAGGATCATGCTGTGAACAAGACATATGACATTGACATGACATACCGACACGACGGGAGT  181
OY      442  CACCCTTCCCGGCATCGCGAGCGCTGTCCATCTCACAGGGAGATGTTTCATCTCTGATG  501
Db      182  CACGACGTTCCGGGCATCGACGGGCTGTCTCCATCTCCAAAGGGGCAACGCTTCACTCGGTG  241
OY      502  TTCAGCCTGTGATTAACCGGAGTCTCTTGATGAGTCAAGCGCTTCAGAGACATCTTG  561
Db      242  TACTCATTTACAGCGCGACAGTCTTGGAGGAGGCTCAAGCCCATCTAACGAACAAATCTGC  301
OY      562  GAGGTCAAGTCTCTGCTGTGAAGAAACAAGACCAAGAGAGGGGCGGAG  606
Db      302  GAGATCAAAAGGAGGACGTGAGCATCTCCCATCATGTGTTGGTGGG  346

RESULT 8
US-08-702-344-26
; Sequence 26, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavalley, Edward
; APPLICANT: Racle, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
;

```



APPLICATION NUMBER: US 08/697,766  
FILING DATE: 29-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragoras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MNI-007CPV2CPA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2184 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 569..1616  
US-08-955-918C-1

Query Match 3.7%; Score 113.8; DB 3; Length 2184;  
Best Local Similarity 89.6%; Pred. No. 2.3e-13;  
Matches 121; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 2948 ACACATATCAATATAAGTGGTTGTTCACAAAAA 3007  
Db 2041 ACATCCCAATATAAGCCGCTTGCGMAAAAAA 2100

Qy 3008 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3067  
Db 2101 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2160

Qy 3068 AAAAAAAAAAAAAA 3082  
Db 2161 AAAAAAAAAAAAAA 2175

RESULT 12  
US-08-697-766A-1  
Sequence 1, Application US/08697766A  
Patent No. 6399760  
GENERAL INFORMATION:  
APPLICANT: Gimeno, Carlos; Kley, Patrick; and Moore, Karen J.  
TITLE OF INVENTION: RP Compositions and Therapeutic and  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,766A  
FILING DATE: 29-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Silveri, Jean M.  
REGISTRATION NUMBER: 39,030  
REFERENCE/DOCKET NUMBER: MNI-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2184 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 569..1616  
US-08-697-766A-1

Query Match 3.7%; Score 113.8; DB 4; Length 2184;  
Best Local Similarity 89.6%; Pred. No. 2.3e-13;  
Matches 121; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 2948 ACACATATCAATATAAGTGGTTGTTCACAAAAA 3007  
Db 2041 ACATCCCAATATAAGCCGCTTGCGMAAAAAA 2100

Qy 3008 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3067  
Db 2101 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2160

Qy 3068 AAAAAAAAAAAAAA 3082  
Db 2161 AAAAAAAAAAAAAA 2175

RESULT 13  
US-09-363-708-3  
Sequence 3, Application US/09363708  
Patent No. 6399747  
GENERAL INFORMATION:  
APPLICANT: Schmandt, et al.  
TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/363,708  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 01017/34451  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6430  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2246 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
DESCRIPTION: /desc = "mouse pAL cDNA"  
US-09-363-708-3

Query Match 3.7%; Score 113.4; DB 4; Length 2246;  
Best Local Similarity 91.6%; Pred. No. 2.8e-13;  
Matches 120; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY	2952	TATCATATAAAGTGGCTTGTTCACAAAAAAAAAAAAAAAAAAAAAAAAA	3011
Db	2112	TATTATATTAAGTGTACTTGACAAAAAAAAAAAAAAAAAAAAAAAAA	2171
QY	3012	AAA	3072
Db	2172	AAA	2231
QY	3072	AAAAAAAAAAAAA 3082	
Db	2232	AAAAAAAAAAAAA 2242	

RESULT 14  
US-09-083-587-3

DESCRIPTION:	/desc = "mouse PAL cDNA"
US-09-083-587-3	

Query Match	3.7%	Score 113.4	DB 4	Length 2246
Best Local Similarity	91.6%	Pred. No. 2.8e-13		
Matches 120; Conservative	0	Mismatches 11	Indels 0	Gaps 0

```

Sequence 50, Application US/09328475C
Patent No. 6476207

GENERAL INFORMATION:
APPLICANT: Zhang, Jimmy
APPLICANT: Astel, Jon H.
APPLICANT: Carroll III, Eddie
APPLICANT: Endege, Wilson O.
APPLICANT: Ford, Donna M.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
APPLICANT: Steinmann, Kathleen E.
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
FILE REFERENCE: 1532_002/200130_463
CURRENT APPLICATION NUMBER: US/09/328,475C
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 341
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 50
LENGTH: 1024
TYPE: DNA
ORGANISM: Homo Sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1024)
OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-50

```

Query Match	3.7%;	Score 113.2;	DB 4;	Length 1024;
Best Local Similarity	92.2%;	Pred. No. 2.4e-13;		
Matches 18;	Conservative	0;	Mismatches 10;	Indels 0;
			Gaps	0

Search completed: June 20, 2004, 08:42:52  
Job time : 215 secs

Genoscope Centre National de Séquençage  
BP 191 91006 EVRY cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
10498.f For more information about this cluster, see  
<http://www.genoscope.cns.fr/>  
[cgi-bin/cluster.cgi?seg=CS0D0C158F05NP1&cluster=10498.f](http://cgi-bin/cluster.cgi?seg=CS0D0C158F05NP1&cluster=10498.f). Contact :  
Peng Liang Email : [liangliang@lscotech.com](mailto:liangliang@lscotech.com) URL :  
<http://fulllength.invitrogen.com>, Invitrogen Corporation 1600





QY 363 CATGAGAGACTTCCACCGTAAGTATACATCCGCGGACATGTATACAGTCTGACAT 422  
 DB 181 CATGAGAGACTTCCACCGTAAGTATACATCCGCGGACATGTATACAGTCTGACAT 240  
 QY 423 CTTGATATCTTGGCAACACCCCTTCCCGGCAATGGAGGCTGTCTCATCTTCAAGG 482  
 DB 241 CTTGATATCTTGGCAACACCCCTTCCCGGCAATGGAGGCTGTCTCATCTTCAAGG 300  
 QY 483 GGATGTCTTCACTCTGGTGTTCAGCTGTGATACCGGAGTCTTCAATGAGTCAAGCG 542  
 DB 301 GGATGTCTTCACTCTGGTGTTCAGCTGTGATACCGGAGTCTTCAATGAGTCAAGCG 360  
 QY 543 CTTTCAGAGAGATCTCTGAGTCAAGTCTGCTGCTGAAGAACAGACAGAGAGCGGC 602  
 DB 361 CTTTCAGAGAGATCTCTGAGTCAAGTCTGCTGCTGAAGAACAGACAGAGAGCGGC 420  
 QY 603 GAGAGTCCCATGCTGATCTGTGCAACAGAGACGACGCGAGCTGTGCGCCAGT 662  
 DB 421 GAGAGTCCCATGCTGATCTGTGCAACAGAGACGACGCGAGCTGTGCGCCAGT 480  
 QY 663 GCGGACACCGAGGCGGAGCTGTGTGTGCGGAGAGAACTCCGCTTACTTCAAGT 722  
 DB 481 GCGGACACCGAGGCGGAGCTGTGTGTGCGGAGAGAACTCCGCTTACTTCAAGT 540  
 QY 723 GTGGGCGAAGAAACCAACGTCGAGATGTTCTAGTCTTCAAGTATGCGCAA 782  
 DB 541 GTGGGCGAAGAAACCAACGTCGAGATGTTCTAGTCTTCAAGTATGCGCAA 600  
 QY 783 GCTGCGACAGAGATGAGCCCGGCTGATCGAAGATCTCGTGCAGTACGCTGACG 842  
 DB 601 GCTGCGACAGAGATGAGCCCGGCTGATCGAAGATCTCGTGCAGTACGCTGACG 660  
 QY 843 CTTTCACACCGAGGCTTGTGATGCGCGGCTGCAAGAGATGAGCGCTTATGCAAGT 902  
 DB 661 CTTTCACACCGAGGCTTGTGATGCGCGGCTGCAAGAGATGAGCGCTTATGCAAGT 720  
 QY 903 CTTGCGCTTGTGCGCGCGCGCGCGCGCTCAAGTCAAGTCAAGTCAAGTCAAGT 962  
 DB 721 CTTGCGCTTGTGCGCGCGCGCGCGCGCGCTCAAGTCAAGTCAAGTCAAGT 780  
 QY 963 CTTTGGGAAAGGCGGCGCGCTGAGAGGAGCAAGTCAAGTCAAGTCAAGTCAAGT 1022  
 DB 781 CTTTGGGAAAGGCGGCGCGCTGAGAGGAGCAAGTCAAGTCAAGTCAAGTCAAGT 840  
 QY 1023 TGGGGCGGCGCTTGGGCAAGTCTTCAAGGAGTGTGCGCGCGCGCTCAAGTCAAGT 1082  
 DB 841 TGGGGCGGCGCTTGGGCAAGTCTTCAAGGAGTGTGCGCGCGCGCTCAAGTCAAGT 900  
 QY 1083 TCCCA-CCGAGGCGCGCGCGCGCGCTTCAAGTCAAGTCAAGTCAAGTCAAGT 1139  
 DB 901 TTTCCACCGGAGCGCGCGCGCGCGCTTGTGCTTCAAGTCAAGTCAAGTCAAGT 960  
 QY 1140 GCGCGCGGCGCGCTG 1153  
 DB 961 GCGCGCGGCGCGCG 974

RESULT 3  
 B0954076 947 bp mRNA linear EST 21-AUG-2002  
 LOCUS AGENCOURT 8866147 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6464135  
 DEFINITION 5', mRNA sequence.  
 ACCESSION B0954076  
 VERSION B0954076.1 GI:22369554  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE JOURNAL  
 Unpublished (1999)

COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
<http://image.llnl.gov>  
 Plate: LINC2656 row: 1 column: 24  
 High quality sequence stop: 668.  
 Location/Qualifiers  
 1..947  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6464135"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_101"  
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

FEATURES  
 source

ORIGIN  
 Query Match 26.5%; Score 817.8; DB 13; Length 947;  
 Best Local Similarity 97.4%; Pred. No. 4.6e-82;  
 Matches 872; Conservative 0; Mismatches 16; Indels 7; Gaps 4;  
 QY 183 CCGGCGCGCGCATTCGCCAGCGCGCATGATGAGATCTTGTCCAGCGGAACTGCAC 242  
 DB 1 CCGGCGCGCGCATTCGCCAGCGCGCATGATGAGATCTTGTCCAGCGGAACTGCAC 60  
 QY 243 GCTCAGTGTGCGCGCGCAAAATCTATCCGCAATGTGTGTCTGTGGTCTCTCGGTTGGG 302  
 DB 61 GCTCAGTGTGCGCGCGCAAAATCTATCCGCAATGTGTGTCTGTGGTCTCTCGGTTGGG 120  
 QY 303 CAAGAGCTCATGCTGTGCTGCTTCCATATGAGCGCGCTTGAAGGACAGTACACACCCAC 362  
 DB 121 CAAGAGCTCATGCTGTGCTGCTTCCATATGAGCGCGCTTGAAGGACAGTACACACCCAC 180  
 QY 363 CATGAGAGACTTCCACCGTAAGTATACATCCGCGGACATGTATACAGTCTGACAT 422  
 DB 181 CATGAGAGACTTCCACCGTAAGTATACATCCGCGGACATGTATACAGTCTGACAT 240  
 QY 423 CTTGATATCTTGGCAACACCCCTTCCCGGCAATGGAGGCTGTCTCATCTTCAAGG 482  
 DB 241 CTTGATATCTTGGCAACACCCCTTCCCGGCAATGGAGGCTGTCTCATCTTCAAGG 300  
 QY 483 GGATGTCTTCACTCTGGTGTTCAGCTGTGATACCGGAGTCTTCAATGAGTCAAGCG 542  
 DB 301 GGATGTCTTCACTCTGGTGTTCAGCTGTGATACCGGAGTCTTCAATGAGTCAAGCG 360  
 QY 543 CTTTCAGAGAGATCTCTGAGTCAAGTCTGCTGCTGAAGAACAGACAGAGAGCGGC 602  
 DB 361 CTTTCAGAGAGATCTCTGAGTCAAGTCTGCTGCTGAAGAACAGACAGAGAGCGGC 420  
 QY 603 GAGAGTCCCATGCTGATCTGTGCAACAGAGACGACGCGAGCTGTGCGCCAGT 662  
 DB 421 GAGAGTCCCATGCTGATCTGTGCAACAGAGACGACGCGAGCTGTGCGCCAGT 480  
 QY 663 GCGGACACCGAGGCGGAGCTGTGTGTGCGGAGAGAACTCCGCTTACTTCAAGT 722  
 DB 481 GCGGACACCGAGGCGGAGCTGTGTGTGCGGAGAGAACTCCGCTTACTTCAAGT 540  
 QY 723 GTGGGCGAAGAAACCAACGTCGAGATGTTCTAGTCTTCAAGTATGCGCAA 782  
 DB 541 GTGGGCGAAGAAACCAACGTCGAGATGTTCTAGTCTTCAAGTATGCGCAA 600

RESULT 5	
B1754083	B1754083 888 bp mRNA linear EST 25-SEP-2001
LOCUS	
DEFINITION	G03027639P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5198024 5' , mRNA sequence.
ACCESSION	B1754083
KEYWORDS	B1754083.1 GI:15745661
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Mammalia; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Eukaryota; Euthera; Primates; Carnathini; Homnidae; Homo. 1 (bases 1 to 888)
REFERENCE	

AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

**ORIGIN**

Query Match	25.4%;	Score 783.4;	DB 12;	Length 888;
Best Local Similarity	96.4%;	Pred. No. 3.2e-78;		
Matches 855;	Conservative 0;	Mismatches 26;	Indels 6;	Gaps 5;

QY	122	CAGCAGGACAGCTCCCGCAGCTCCCGGCGCTTCCAGGACAGCTCTGTAGCCGTCAGAG	181
Db	1	CAGCAGGACAGCTCCCGCAGCTCCCGGCGCTTCCAGGACAGCTCTGTAGCCGTCAGAG	60
QY	182	GCCCGGCCCCCATTCACGCCCCGAGCCATGATGAGACTTTGTCCAGCGGAACTTCA	241
Db	61	GCCCGGCCCCCATTCACGCCCCGAGCCATGATGAGACTTTGTCCAGCGGAACTTCA	120
QY	242	CGCTCAGTGGCCCCGCAAAAATCTATACCGCATGTGTCTGTGGGTGCGCTCGGGTGG	301
Db	121	CGCTCAGTGGCCCCGCAAAAATCTATACCGCATGTGTGTGTGTGGTGTCTTCTGGGTGG	180
QY	302	GCAAGAGCTCATCTGTGTGCGCTTCTCAATGGCGCGTTTGAAGACCAAGTACACCCA	361
Db	181	GCAAGAGCTCATCTGTGTGCGCTTCTCAATGGCGCGTTTGAAGACCAAGTACACCCA	240
QY	362	CCATGAGGACCTTCCACCGTAGGATATACATCCCGGGGGGACATGTACCAAGTGGACA	421
Db	241	CCATGAGGACCTTCCACCGTAGGATATACATCCCGGGGGGACATGTACCAAGTGGACA	300
QY	422	TCTGTGATACCTCTGTGCAACCAACCCCTTCCCGGCATGCGAGGCTGTCACTCTCAAG	481
Db	301	TCTGTGATACCTCTGTGCAACCAACCCCTTCCCGGCATGCGAGGCTGTCACTCTCAAG	360
QY	482	GGGATCTCTTACCTCGTGTGTCAACCGTGGATTAACCGGGAATCTTGTGATGAGGTCAAGC	541
Db	361	GGGATCTCTTACCTCGTGTGTCAACCGTGGATTAACCGGGAATCTTGTGATGAGGTCAAGC	420
QY	542	GCGTTCAGAGCAGATCTTGAGAGTCAAGTCTTGCTTAAAGAACAGACCAAGAGGCGG	601
Db	421	GCGTTCAGAGCAGATCTTGAGAGTCAAGTCTTGCTTAAAGAACAGACCAAGAGGCGG	480
QY	602	CGGAGCTGCCCATGTGTCACTGTGTGCAACAAAGAACACAGGAGAGTGTGCGCGTCAAG	661
Db	481	CGGAGCTGCCCATGTGTCACTGTGTGCAACAAAGAACACAGGAGAGTGTGCGCGTCAAG	540

[illegible]

RESULT 6	BO954352	918 bp	mRNA	linear	EST 21-NUG-2002
LOCUS	AGNCOURT_8794894 NIH_MGC_18	Homio sapiens	CDNA clone	IMAGE:6374730	
DEFINITION	5', mRNA sequence.				
ACCESSION	BO954352				
VERSION	BO954352.1	GI:22369830			
KEYWORDS	EST.				
SOURCE	Homio sapiens (human)				
ORGANISM	Homio sapiens				
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 918)				
JOURNAL	NIH-MGC <a href="http://mgi.nci.nih.gov/">http://mgi.nci.nih.gov/</a> .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a>				
	Tissue Procurement: DCTD/DTF/Gazdar				
	CDNA Library Preparation: Rubin Laboratory				
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	<a href="http://image.lnl.gov">http://image.lnl.gov</a>				
	Plate: LCM2554 row: c column: 19				
	High quality sequence, stop: 578.				

**FEATURES**  
**source**

ORIGIN	-	-
Query Match	24.9%	Score 766.8; DB 13; Length 918
Best Local Similarity	98.3%;	Pred. No. 2.2e-76;

Matches	828; Conservative	0; Mismatches	7; Indels	7; Gaps	5;
Qy	1017	GGATGCTGGGGGCGGGGCTTTGGCCAGTGGCTTCAGGGAGTGGCCCCAATGCCACTGTG	1076		
Db	1	GGAATGCTGGGGGCGGGGCTTTGGCCAGTGGCTTCAGGGAGTGGCCCCAATGCCACTGTG	60		
Qy	1077	CGCATCTCCCAACGAGGCGCCCGGAGGAGTCTTTTCAACAGACTTTAGGACAGCAGACTG	1136		
Db	61	CGCATCTCCCAACGAGGCGCCCGGAGGAGTCTTTTCAACAGACTTTAGGACAGCAGACTG	120		
Qy	1137	GAGGCCCCCGGGCGCTGGCTTCGACATTTGCTGCTTCACAGCTTTCTGAGTCC	1196		
Db	121	GAGGCCCCCGGGCGCTGGCTTCGACATTTGCTGCTTCACAGCTTTCTGAGTCC	180		
Qy	1197	GCTTGTCCACAGCTCTTGGTGTGTTTATCTCTCTTGGGAGAGACATCTTGCACCC	1256		
Db	181	GCTTGTCCACAGCTCTTGGTGTGTTTATCTCTCTTGGGAGAGACATCTTGCACCC	240		
Qy	1257	TCAAGAGTTAGGACAGACTCAAGTTACACTTCTCTCTGGGGTTGGAAAGAAATGTTG	1316		
Db	241	TCAAGAGTTAGGACAGACTCAAGTTACACTTCTCTCTGGGGTTGGAAAGAAATGTTG	300		
Qy	1317	ATGCCAGAGGGGTGAGATTTGCTGCTCATATGAGCTCTCTGGAGCAAGCTCAGATG	1376		
Db	301	ATGCCAGAGGGGTGAGATTTGCTGCTCATATGAGCTCTCTGGAGCAAGCTCAGATG	360		
Qy	1377	AAAAGACACAGAGGGCCAGATGAGAAAGTCTCTCTCTCTGGCATTAACCCAGCTT	1436		
Db	361	AAAAGACACAGAGGGCCAGATGAGAAAGTCTCTCTCTCTGGCATTAACCCAGCTT	420		
Qy	1437	GGTTGGGTGGGAGCTGGGAGAACTTCTCCCAAGCCCTGCAACTTACGCTGTGTTG	1496		
Db	421	GGTTGGGTGGGAGCTGGGAGAACTTCTCTCCCAAGCCCTGCAACTTACGCTGTGTTG	480		
Qy	1497	AGTGCCTCTGTCACCCCTCTCCACCCCGACACACACAAAGTTGGCCCGAGCTGCGC	1556		
Db	481	AGTGCCTCTGTCACCCCTCTCCACCCCGACACACACAAAGTTGGCCCGAGCTGCGC	540		
Qy	1557	TGACATTGAGCCAGTGGAGTCTGTGTCTGAA--GGGGGCTGGCCACACTTCTTACCA	1614		
Db	541	TGACATTGAGCCAGTGGAGTCTGTGTCTGAA--GGGGGCTGGCCACACTTCTTACCA	600		
Qy	1615	CGCCACCACTTGAACACAGCCACCTCTGACCGGCTTCTCAGACTTCTCTCTAGGT	1674		
Db	601	CGCCACCACTTGAACACAGCCACCTCTGACCGGCTTCTCAGACTTCTCTCTAGGT	660		
Qy	1675	CCCTCCGCGCCGACAGTT--GTGCTTTGTGTGGTTGCAAGCTGTTTTGTGTCATGTAAGT	1733		
Db	661	CCCTCCGCGCCGACAGTT--GTGCTTTGTGTGGTTGCAAGCTGTTTTGTGTCATGTAAGT	720		
Qy	1734	AGTAAATATGAAATCATTTGTACTGTAAAGGCTTGTGATCTCC--TCTTGGCCAGGCGC	1792		
Db	721	AGTAAATATGAAATCATTTGTACTGTAAAGGCTTGTGATCTCC--TCTTGGCCAGGCGC	780		
Qy	1793	TCACCCAGTTCAGATTCACAGGCG--TCACCCGCGGAGC--CTTCTCTCTCTCTCCAA	1849		
Db	781	TCACCCAGTTCAGATTCACAGGCGCTTCACCCGCGGAGCGCTTCTCTCTCTCTCCAA	840		
Qy	1850	CA 1851			
Db	841	AA 842			

RESULT 7  
 LOCUS BX327655 891 bp mRNA linear EST 01-MAY-2003  
 DEFINITION BX327655 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 CDNA clone CS0DC015YK10 5-PRIME, mRNA sequence.  
 ACCESSION BX327655  
 VERSION BX327655.1 GI:30307667  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE	1 (bases 1 to 891)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10498.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAG0072D04_C 500604.1cluster=10498.f. Contact : Feng Liang Email : fliang@life-tech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0BAG0072D04_C500604.1.
FEATURES	Location/Qualifiers
source	1..891 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DC015YK10" /cissue="NEUROBLASTOMA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	
Query Match	22.6%; Score 696.8; DB 13; Length 891;
Best Local Similarity	96.6%; Pred. No. 1.4e-68;
Matches	744; Conservative 0; Mismatches 22; Indels 4; Gaps 3;
Qy	26 GCCGCGCCACCGCGGCTCCCGACAGCGAGGAGATCCCGCGAGTGAACCCGGG 85
Db	121 GCCGCGCCACCGCGGCTCCCGACAGCGAGGAGATCCCGCGAGTGAACCCGGG 180
Qy	86 AGCCACACAGACTCTGGGAGGCTCGCGGCTGAGACAGCGAGTCTCCCGAGCTCC 145
Db	181 AGCCACACAGACTCTGGGAGGCTCGCGGCTGAGACAGCGAGTCTCCCGAGCTCC 240
Qy	146 CGGCGCTTCAGGAGCTCTGAGCGGTGACAGAGCGCGGCGCGCATTCGAGCCG 205
Db	241 CGGCGCTTCAGGAGCTCTGAGCGGTGACAGAGCGCGGCGCGCATTCGAGCCG 300
Qy	206 GAGCATGATGAAGACTTGTTCAGCGGGAATGACAGCTCAGTGTGCGCCGCAAACT 265
Db	301 GAGCATGATGAAGACTTGTTCAGCGGGAATGACAGCTCAGTGTGCGCCGCAAACT 360
Qy	266 CATACCGATGTGTGTGCTGCGGCTCTCGGCGTGGGAAAGTCCATCTGTCTGCT 325
Db	361 CATACCGATGTGTGTGCTGCGGCTCTCGGCGTGGGAAAGTCCATCTGTCTGCT 420
Qy	326 TCTTCAATGGCGCTTGAAGACAGTACACACCACCATGAGAGACTTCCACGGTAA 385
Db	421 TCTTCAATGGCGCTTGAAGACAGTACACACCACCATGAGAGACTTCCACGGTAA 480
Qy	386 TATACACATCCGCGGACAGTATGACAGCTGACATCTCGATGATCTTGGCAACAC 445
Db	481 TATACACATCCGCGGACAGTATGACAGCTGACATCTCGATGATCTTGGCAACAC 540
Qy	446 CTTTCCCGGCAATGGCGAGCTGTTCATCTGACAGGGAATGTCTTATCTCTGTGTTCA 505
Db	541 CTTTCCCGGCAATGGCGAGCTGTTCATCTGACAGGGAATGTCTTATCTCTGTGTTCA 600
Qy	506 GCTCGAATAACGGGAGTCTTTCATGAGTCAAGCGCTCAGAAAGAGATCTCGAGG 565
Db	601 GCTCGAATAACGGGAGTCTTTCATGAGTCAAGCGCTCAGAAAGAGATCTCGAGG 660
Qy	566 TCAAGTCTGCTGCTGAAGAACCAAGCAAGAGAGCGG--CGAGAGTGGCCCATGATCTGT 624

Db	Accession	Source	Organism	Reference
Db	661	TCAGTCTGCTTGAAGAACAGACCAAGAGGCGGACGAGCTGCTCATGTCTCT		
Qy	625	GGCAACAAAGAACGACCAAGGAGCTGTGCGGACGAGTCCACCAAGAGCGAGCTG		
Db	721	GGCAACAAAGAACGACCAAGGAGCTGTGCGGACGAGTCCACCAAGAGCGAGCTG		
Qy	685	CTGTGTCTGGGCGACGAGAACTCCGCTTACTTTCGAGTGTGGGCGAAGAACCAAC		
Db	781	CTGTGTCTGGGCGACGAGAACTCCGCTTACTTTCGAGTGTGGGCGAAGAACCAAC		
Qy	745	GTGACGAGATGTTTACGTGCTTTCAGATGAGCAATGCGCAAGCTGCCACAGA		
Db	839	GTGACGAGATGTTTACGTGCTTTCAGATGAGCAATGCGCAAGCTGCCACAGA		
RESULT 8	LOCUS	AK015898	1300 bp	mRNA
DEFINITION				linear
ACCESSION		AK015898		HTC 20-SEP-2003
VERSION		AK015898.1		
KEYWORDS		HTC; CAP trapper.		
SOURCE		Mus musculus (house mouse)		
ORGANISM		Mus musculus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE		Carninci, P. and Hayashizaki, Y.		
JOURNAL		High-efficiency full-length cDNA cloning		
MEDLINE		Meth. Enzymol. 303, 19-44 (1999)		
PUBMED		99279253		
REFERENCE		10349636		
TITLE		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
JOURNAL		Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
MEDLINE		Normalization and subtraction of cap-trapper-selected cDNAs to		
PUBMED		prepare full-length cDNA libraries for rapid discovery of new genes		
REFERENCE		Genome Res. 10 (10), 1617-1630 (2000)		
TITLE		20499374		
JOURNAL		11042159		
MEDLINE		3		
PUBMED		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, T., Sasaki, N., Carninci, P.,		
REFERENCE		Komo, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,		
AUTHORS		Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,		
TITLE		Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
JOURNAL		Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,		
MEDLINE		Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J.,		
PUBMED		Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
REFERENCE		RIKEN integrated sequence analysis (RISA) system--384-format		
TITLE		sequencing pipeline with 384 multichannel sequencer		
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE		20530913		
PUBMED		11076861		
REFERENCE		4		
AUTHORS		The RIKEN Genome Exploration Research Group Phase II Team and the		
TITLE		FANTOM Consortium.		
JOURNAL		Functional annotation of a full-length mouse cDNA collection		
MEDLINE		Nature 409, 685-690 (2001)		
PUBMED		15		
REFERENCE		The FANTOM Consortium and the RIKEN Genome Exploration Research		
AUTHORS		Group Phase I & II Team.		
TITLE		Analysis of the mouse transcriptome based on functional annotation		
JOURNAL		of 60,770 full-length cDNAs		
MEDLINE		Nature 420, 563-573 (2002)		
PUBMED		6 (bases 1 to 1300)		
REFERENCE		Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,		
AUTHORS		Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,		
TITLE		Funano, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,		
JOURNAL		Hitaoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,		
MEDLINE		Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komo, H., Kouda, M.,		
PUBMED				
REFERENCE				
AUTHORS				

TITLE  
JOURNAL

COMMENT

CDS

FEATURES  
source

ORIGIN

Query Match      22.5%; Score 693.2; DB 11; Length 1300;  
Best Local Similarity    87.0%; Pred. No. 2.9e-66;  
Matches    774; Conservative    0; Mismatches 113; Indels    3; Gaps    1;

200 AGCCCGAGCAGCATGATGAAGCTTTGGCCAGCGGAACTGCAGCGTCAGTGTGCCGCA 259  
Db AGCCAGGACATATAGAAGCTTGTCAGTGGAACTGCACACTAATGTGCTGCTA 424

260 AAATCTATAACCGATGATGATGCTGCGATGCTTCGGATGGGCAAGAGCTCATCTGT 319  
Db AAGACTCTTAACCGATGATGATGCTGCGATGCTTCGGATGGGCAAGAGCTCATCTGT 484

320 CTGCGTCTCATATGCGCGCTTTGAGACAAGTACACCCAGCATGAGAGCTTCCACC 379

Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,  
Munura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okio,T., Owa,C.,  
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D.,  
Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y.,  
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,  
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,  
Yoshino,M., Muramatsu,M. and Hayaishizaki,Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayaishizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Science Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5'-GAGAGGAGGAGATCCAGACGCTCTTTTTTTTTTTVN 3'], cDNA was  
prepared by using trehalase thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. Second  
strand cDNA was prepared with the primer adapter of sequence [5'  
GAGAGGAGGATTCGACGTTAATTAAATATCCCCCCCCC 3']. cDNA was cleaved  
with BamHI and XhoI. cDNA of size comprised longer than 7 kb was  
selected before cloning. Vector: a modified pluscript KS(+) after  
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'  
end: BamHI. Host: DH10B.

Location/Qualifiers

1..1300  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strains="C57BL/6J"  
/db\_xref="PANTOM,DB:4930526B11"  
/db\_xref="MG1:1908452"  
/db\_xref="taxon:100900"  
/clone="4930526B11"  
/sex="male"  
/tissue\_type="testis"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
376..1176  
/note="unnamed protein product; putative  
weakly similar to DEXRA51 [Rattus norvegicus]  
(SPR109JKF8, evidence: FASTV, 60.5%id, 98.2%length,  
match=798)"  
/codon\_start=1  
/protein\_id="DAB30023.1"  
/db\_xref="GI:12854417"  
/translation="MMKTLSSGNLTAVPAKNRYMVVLASRVGSISYRFNGRR  
EDQPTIEDPHRKRVNIHEDMTQLDLIDTSGNHPFAMRRLSLTDGVFLIVSLNS  
RESDEVAERLQKOLLKVNSCLKNKTKEAALPMVT CANKNDHSLLCRQVPAMEELLV  
SGSENCAYFEVSAKNTNVNEMFYLVLSMAKLPLEMSPALHHKISTVGDAFFRPFC  
MRRTKVAGAMVSPFARRBSVNSDLKYIRAKVILREGARERDRCSIO"

[illegible]

```

clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/INTL at:
http://image.llnl.gov
Plate: LIMS996 row: C column: 08
High quality sequence step: 688.
Location/Qualifiers
1. .730
FEATURES
source
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone IMAGE:4358335"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_84"
/note="Organ: adrenal gland; Vector: pCMV-Sport6; Site_1:
NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT
primed. Average insert size 1,229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

## ORIGIN

Query Match	22.2%	Score 684.6;	DB 10;	length 730;
Best Local Similarity	97.7%	Pred. No. 3,4e-67;		
Matches 716;	Conservative	0;	Mismatches 14;	Indels 3;
				Gaps 2;

QY 1712 CTGTTTCGTGCATGTATAGTAGTAGAAGTGAATCATTTGTACTGTAAAAAGCCTAGTG 1771

Db 1 C T G T T T C G T C A T G T A T A G T A G T A G A A A T G S A A T C A T T G T A C T G T A A A A G C C T A G T G 60

1772 ACTCCCTCCTTGGCCAGGCGCTCACCAGTTGAGATCCACGGGCTTCACCGGGACGGCT 1831

Db 61 ACTICCTCCTTGCCAGGCCCCACCCAGTTCAGATCCACGCTCCACCCGGACGCC 120

1832 TCTCTCTCTGCTCCCAACAGGGTTCGGTGGCCGTTCTGATCAGACATTGATCCCG 103

[illegible][illegible]

2011

[illegible]

2012 AGCAGGACCCCTTTCTCCCTTACGAAATTACGGTACAGCTCACCTTGAACCAAG 2071

301 AGCAACCAAGCCCCACTTTTGTCTCCATGCAAAATTCAGGGTACAGCTCCACCTTACACCAG 360

2072 GCTGCCCTCTACTGCTGCTCGTTCTCAAGCAATTATTAGCACCTACTGGTGCTGGCTT 2131

Db 361 GCGCCCTTACTGTGCTCGTTCTCAAGATTATTAGACACTAGGGTGGGTT 420

2132 CACTGTGTCTTAGAACAACAAGGGTCCCAAGTCTGGCTTGCCCGCCCTGTGCC 219

Db 421 CACGTGTCTTAGGAACCAAGAGGCTCCCACTCTGGCTTCTGCCCGCCCTGCTGCC 480

2192 CCACACCTTCTGCAACACAGCGGTGGGAGGCCGGGAGGAGCACTGGGACCCAGAAC 225

481 CCACCACCTTCTGCACACACAGCGTGGGAGCGCGGAGAGCAGCTGGACCCAGAAC 540

2252 TGAGCCTGGAGGGATCCGACAGAAAGCTCAGGGCGGGTCTTCTCTTGTGCCCGGGAT 2311

Db 541 TGAGCTGGGAGGGCTCCGACAGAAAGCTCAGGGCGGGTCTTCTCTTGTGCCCCGGGAT 600

2312 TGGGCTATGCTGGGTCACCATGTACTCAGGCATGGTGGGTTTGAACCCATAAACCA 2317

601 TGGGCTATGCTGGGTACCAACCATGTACTCAGGCATGGTGGGTTT--GAACCAAAACCAA 658

2372 AGGCCCTGTCTACGCTCTTAACAGATATATTGATTTAAACICCTAAACATATT 243

Db 659 AGGCCCTGTGATCAGCTCTTACAGTAATA-TAAGTATCAATCTTCCCTACCAATAT /1

QY	2432	GAAGTTTAGGAC	2444
Db	718	GAAGTTTAGGAC	730
RESULT 10			
LOCUS	BG119387	1185 bp	mRNA linear EST 30-JAN-2001
DEFINITION	6023490701 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:444168 5',		
ACCESSION	BG119387		
VERSION	BG119387.1	GI:12612893	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 1185)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strusberg, Ph.D.		
	Email: cgsbbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: L1AM10219 row: k column: 17		
	High quality sequence stop: 679.		
FEATURES			
source	Location/Qualifiers		
	1..1185		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:444168"		
	/tissue_type="adenocarcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_id="NIH MGC 90"		
	/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;		
	Site 2: SalI; Cloned unidirectionally; oligo-dT primed.		
	Average insert size 1.7 kb. Library enriched for		
	full-length clones and constructed by Life Technologies.		
	Note: this is a NIH_MGC Library."		
ORIGIN			
	Query Match	21.5%;	Score 661.8; DB 12; Length 1185;
	Best Local Similarity	88.8%;	Pred. No. 9.5e-65;
	Matches 775; Conservative	0; Mismatches 87; Indels 11; Gaps 5;	
QY	1904	ACGGTTTACACAACTTGCACAGCGTGGGGGAGGCGACGACCTGCTTTTAT	19633
Db	1	ACGGTTTACACAACTTGCACAGCGTGGGGGAGGCGACGACCTGCTTTTAT	60
QY	1964	GCTCCCAATTTGACAGAGATACCAACGAGCTCGAGGGGAGACGATGAGACAGGCCC	20233
Db	61	GCTCCCAATTTGACAGAGATACCAACGAGCTCGAGGGGAGACGATGAGACAGGCCC	120
QY	2024	CACCTTTGTCCTCGACAAATTCAGGGGTACAGCTTCACCTAGAACCAAGGCTGCTTAC	20833
Db	121	CACCTTTGTCCTCGACAAATTCAGGGGTACAGCTTCACCTAGAACCAAGGCTGCTTAC	180
QY	2084	TGTGCTGTTCTCAAGACTTATTAACACCTACTGGGTGCTGCTTCACTGTGTCTTA	21433
Db	181	TGTGCTGTTCTCAAGACTTATTAACACCTACTGGGTGCTGCTTCACTGTGTCTTA	240
QY	2144	GGAACCAAGAGGTCCTCCAGTCTGCTGCTGCGCCGCTGCTGCGCCCAACACTTCT	22033
Db	241	GGAACCAAGAGGTCCTCCAGTCTGCTGCTGCGCCGCTGCTGCGCCCAACACTTCT	300
QY	2204	GCACACACAGCGGTGGGAGGCGGGGAGAGACGCTGGGACCCAGAACTGAGCTGGAG	22633

Db	Accession	Source	Length	EST	EST Date
Db	301	GCACACACGCGGTGGGAGGGCGGGAGAGACACTGGGACCAGAACTGAGCTTGGAG	360		
Qy	2264	GGATTCGACAGAAAAGCTCAGGGCGGGCTCTTCCTTGTGCGGGAGATTGGGCTATGCTG	2323		
Db	361	GGATTCGACAGAAAAGCTCAGGGCGGGCTCTTCCTTGTGCGGGAGATTGGGCTATGCTG	420		
Qy	2334	GGTACACCACTATCTCAGGGCATGTGGGTTTGAACCCATMAACCAAGGCCCTTGTCA	2385		
Db	421	GGTACACCACTATCTCAGGGCATGTGGGTTTGAACCCATMAACCAAGGCCCTTGTCA	480		
Qy	2384	TCAGCTCTTAACAAGTATATTTGATTTTAATCTCTTAACAATATTGAAGTTTAGGG	2443		
Db	481	TCAGCTCTTAACAAGTATATTTGATTTTAATCTCTTAACAATATTGAAGTTTAGGG	540		
Qy	2444	CCCTAAGAACCTTAGTATCTTCTATTTGGGCTCTTCTGAGTTTCAG-AGAGGGTAAAGTA	2502		
Db	541	CCCTAAGAACCTTAGTATCTTCTATTTGGGCTCTTCTGAGTTTCAGAGAGGGTAAAGTA	600		
Qy	2503	ACTTCCTTCAGGTCAACAGCAAGTCTGTGGGTGGCAG-AGCAAGCTAGCGCTGGGCAT	2561		
Db	601	ACTTCCTTCAGGTCAACAGCAAGTCTGTGGGTGGCAGAGAAAGCTAGCGCTGGGCAT	660		
Qy	2562	TCAGTACATACCAAGATGTGTCC---TCCTTAATGCTTGGCCCTTGGGGCCTTCAG	2617		
Db	661	TCAGTACATACCAAGATGTGTCC---TCCTTAATGCTTGGCCCTTGGGGCCTTCAG	720		
Qy	2618	GGCTTTGGGACATCTTGTCCTC---AACCCTCCCTAGATCAAGTCTGTAGGGTCCCT	2673		
Db	721	GGCTTTGGGACATCTTGTCCTC---AACCCTCCCTAGATCAAGTCTGTAGGGTCCCT	780		
Qy	2674	GTAGATATTGTGTACAC-CATGCCCATGTATATACAAAGTACACACAGATGTACACAGA	2732		
Db	781	GACATTTGTGTGTACACAGCTGCCAGATGTTATATACAGTCTCCACAGAGTTAACCGA	840		
Qy	2733	TGTACACATGTCTCCAGCCCGCAGCTCTGCATACC	2765		
Db	841	GTGTGCACACAGTCTCCAGCCCGAGATGTGTACCC	873		
RESULT 11					
LOCUS	BE408863	809 bp	mRNA	linear	EST 21-JUL-2000
DEFINITION	601303729F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3638118 5'				
ACCESSION	BE408863				
VERSION	BE408863.1	GI:9345313			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 809)				
TITLE	NIH-MGC htcp://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-remail.nih.gov				
	Tissue Procurement: ATCC				
	cDNA Library Preparation: Ling Hong/Rubin Laboratory				
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.nsl.gov				
	Plate: L1CM38 row: 5 column: 07				
	High quality sequence start: 7				
	High quality sequence stop: 618.				
FEATURES					
source	1..809				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:3638118"				



## ORIGIN

/tissue\_type="choriocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NH MGC 21"  
 /note="Organ: placenta; Vector: pOTB7, Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

Query Match 20.4%; Score 629.8; DB 10; Length 809;  
 Best Local Similarity 97.5%; Pred. No. 4e-61;  
 Matches 693; Conservative 0; Mismatches 12; Indels 6; Gaps 5;

26 GCCGCGCCAGCCGCGGCTCCGACGACGCGGAGAGATCCCGCGAGTACCGGAG 85  
 4 GTCCGCGCCAGCCGCGGCTCCGACGACGCGGAGAGATCCCGCGAGTACCGGAG 63

86 AGCCACACAGACTCTGCGAGGCTCGCGGCTGAGACAGACGACGACTCCCGACGCTCC 145  
 64 AGCCACACAGACTCTGCGAGGCTCGCGGCTGAGACAGACGACGACTCCCGACGCTCC 123

146 CGCGGCTTCGAGGAGCTCTCTGAGCCCTGCGAGGCGCGCGCGCCATTCGACGCCCC 205  
 124 CGCGGCTTCGAGGAGCTCTCTGAGCCCTGCGAGGCGCGCGCGCCATTCGACGCCCC 183

206 GAGCCATGATGAGACTTGTCTCAGCGGGAATCTGACGCTAGTGTGCCCGCAAAAAT 265  
 184 GAGCCATGATGAGACTTGTCTCAGCGGGAATCTGACGCTAGTGTGCCCGCAAAAAT 243

266 CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325  
 244 CATACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303

326 TCCTCAATGCGCGCTTTGAGAGCCAGTACACACCAACCATCGAGAGCTTCCACCTTAAG 385  
 304 TCCTCAATGCGCGCTTTGAGAGCCAGTACACACCAACCATCGAGAGCTTCCACCTTAAG 363

386 TATACAAATCCGCGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 445  
 364 TATACAAATCCGCGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 423

446 CTTTCCCGCGCATGCGAGGCTGTCCATCTTCAACAGGGAGTCTTCACTCTGATGTTCA 505  
 424 CTTTCCCGCGCATGCGAGGCTGTCCATCTTCAACAGGGAGTCTTCACTCTGATGTTCA 483

506 GCTTGGATTAACCGGAGATCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565  
 484 GCTTGGATTAACCGGAGATCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543

566 TCAAGTCTGCGCTTAAGAAACAAGACCAAGAGCGCGAGTGTCCATGATGATGATGATGATGATGAT 624  
 544 TCAAGTCTGCGCTTAAGAAACAAGACCAAGAGCGCGAGTGTCCATGATGATGATGATGATGATGAT 603

625 GGCACCAAGAAAGACACGCGCGAGCT-GTCCGCGCAGAGTGTCCACACCGAGCGCGAGCT 683  
 604 GGCACCAAGAAAGACACGCGCGAGCTGTGATGCGCGAGAGTGTCC-ACAACAGAGCGCGA-CT 660

664 GCTGTGTGTGCGCGACAGAAACT-GGCTTAATCCAGAGTGTGCGCAAAA 710  
 661 GCTGTGTGTGCGCGACAGAAACT-GGCTTAATCCAGAGTGTGCGCAAAA 710

RESULT 12  
 BQ017479/c 629 bp mRNA linear EST 17-JUN-2002  
 LOCUS BQ017479  
 DEFINITION UI-H-DT1-swe-j-21-0-UI.s1 NCI CGAP DT1 Homo sapiens cDNA clone  
 IMAGE:5888372 3', mRNA sequence.  
 ACCESSION BQ017479  
 VERSION BQ017479.1 GI:19752756

## KEYWORDS

## SOURCE

EST.  
 Homo sapiens (human)

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 629)

## AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

## TITLE

Unpublished (1997)

## JOURNAL

Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

## COMMENT

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
 The following repetitive elements were found in this cDNA  
 sequence: 458-519, >MIR#SINE/MIR  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

## FEATURES

## source

Location/Qualifiers  
 1..629  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5888372"  
 /tissue\_type="Metastatic Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP DT1"

/note="Organ: Lung; Vector: pRT3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI CGAP DT1 is a normalized cDNA library containing the  
 following tissue(s): Metastatic Chondrosarcoma in Lung.  
 The library was constructed according to Bonaldo, Lennon  
 and Soares, Genome Research, 6:791-806, 1996. First strand  
 cDNA synthesis was primed with an oligo-dT primer  
 containing a Not I site. Double stranded cDNA was ligated  
 to an EcoR I adaptor, digested with Not I, and cloned  
 directionally into pRT3-Pac vector. The oligonucleotide  
 used to prime the synthesis of first-strand cDNA contains  
 a library tag sequence that is located between the Not I  
 site and the (dT)18 tail. The sequence tag for this  
 library is AACTGTTCCG.  
 TAG TISSUE=Lung metastatic chondrosarcoma  
 TAG LIB=UI-H-DT1  
 TAG\_SEQ=AACTGTTCCG"

## ORIGIN

Query Match 20.3%; Score 625.8; DB 12; Length 629;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-60;  
 Matches 627; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2370 AAAGCCCTTTCATCAGCTCTTAACAGTATTTTATTTAATCTCTAAACATA 2429  
 629 AAAGCCCTTTCATCAGCTCTTAACAGTATTTTATTTAATCTCTAAACATA 570

2430 TTGAAGTTTGAAGGCGCTTAAGAACTTATGATCTTTTGAAGTTCA 2489  
 569 TTGAAGTTTGAAGGCGCTTAAGAACTTATGATCTTTTGAAGTTCA 510

2490 GAGAGGTAAGTAATCTTCTCAGGTCACACAGAGTCTGTGGGTGCGAAGCAGCT 2549  
 509 GAGAGGTAAGTAATCTTCTCAGGTCACACAGAGTCTGTGGGTGCGAAGCAGCT 450

2550 AGCGCTGGCATTAAGTATCAATACACAGATGCTCCCTCTTATGCTTGGCCCTGGG 2609  
 449 AGCGCTGGCATTAAGTATCAATACACAGATGCTCCCTCTTATGCTTGGCCCTGGG 390

2610 GCCTTCAGGCGCTTGAAGATCTTGTCTTCAACCTCTCTCAATCAGTCTGTGAGGGT 2669



Db 389 GCCTTCAGAGGCTTGGAGCATCTTGTCTCAACCTCTCCTAGATCAGTGTGAGGGT 330  
 Qy 2670 CCTGTGATATTTGTGTATACCATGCCCCATGTATATACATACACAGATGTACAC 2729  
 Db 329 CCTGTGATATTTGTGTATACCATGCCCCATGTATATACATACACAGATGTACAC 270  
 Qy 2730 AGATGTACACATGTCTCCAGCCCCAGCTGTGATACCTGTGACACCCAGCTTGGCC 2789  
 Db 269 AGATGTACACATGTCTCCAGCCCCAGCTGTGATACCTGTGACACCCAGCTTGGCC 210  
 Qy 2790 CCTGCTGCTGTGTGTCTCAAGAGCAGCTTCAACCTGTGCTGTCTCCCTCCAGC 2849  
 Db 209 CCTGCTGCTGTGTGTCTCAAGAGCAGCTTCAACCTGTGCTGTCTCCCTCCAGC 150  
 Qy 2850 CACTGCTGAGCTTTCTGAGCAGACAGTACCTTGTGCTGACCGGTGTGCGCCGCTC 2909  
 Db 149 CACTGCTGAGCTTTCTGAGCAGACAGTACCTTGTGCTGACCGGTGTGCGCCGCTC 90  
 Qy 2910 TCACCCAGGACAGCCCGCCAGCAGTATCGGTGTGACCTATCAATAAAGTGGGT 2969  
 Db 89 TCACCCAGGACAGCCCGCCAGCAGTATCGGTGTGACCTATCAATAAAGTGGGT 30  
 Qy 2970 TGTTCACAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2998  
 Db 29 TGTTCACAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1

RESULT 13  
 BE390509 627 bp mRNA linear EST 21-JUL-2000  
 LOCUS 601284127F1 NIH\_MGC\_44 Homo sapiens cDNA clone IMAGE:3605766 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE390509  
 VERSION BE390509.1 GI:9335874  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 627)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LCM254 row: f column: 07  
 High quality sequence stop: 614.  
 Location/Qualifiers  
 1..627  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3605766"  
 /tissue\_type="endometrium, adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_44"  
 /note="Organ: uterus; Vector: pOT7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dt priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G) Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-CDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN  
 Query Match 19.1%; Score 587.4; DB 10; Length 627;

Best Local Similarity 99.4%; Pred. No. 2,3e-56;  
 Matches 621; Conservative 0; Mismatches 1; Indels 3; Gaps 3;  
 Qy 26 GCCGCCGCCAGCCCGCGTCCCGAGCAGCGAGGGAGATCCCGCGCAGTACCCGGG 85  
 Db 3 GCCGCCGCCAGCCCGCGGT-CCGAGCAGCCAGGGGAGATCCCGCGCAGTACCCGGG 61  
 Qy 86 AGCCACACA-GACTCTGGAGAGGCTCGCGCTGAGACAGAGGACCTCCCGCAGTCC 144  
 Db 62 AGCCACACATGACTCTGGAGAGGCTCGCGCTGAGACAGAGGACCTCCCGCAGTCC 121  
 Qy 145 CCGCGCTTCCAGGACAGCTCTGTAGCGGTGACAGAGCCCGCGCCATTCCAGGCC 204  
 Db 122 CCGCGCTTCCAGGACAGCTCTGTAGCGGTGACAGAGCCCGCGCCATTCCAGGCC 181  
 Qy 205 CGAGCCATGATGAAGACTTTGTCCAGCGGGAATGACAGCTCACTGTGCGCCAAAAC 264  
 Db 182 CGAGCCATGATGAAGACTTTGTCCAGCGGGAATGACAGCTCACTGTGCGCCAAAAC 241  
 Qy 265 TCATACCGCATGTGTGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 324  
 Db 242 TCATACCGCATGTGTGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 301  
 Qy 325 TTCTCATGTGGCGCTTGTGAGAGCA-GTACACACCCACCATGAGACTTCCACCGTAA 383  
 Db 302 TTCTCATGTGGCGCTTGTGAGAGCATGTACACACCATGAGACTTCCACCGTAA 361  
 Qy 384 GTATACACATTCGCGCGGCGCATGTACACAGCTTGACATCTGTGCAACCA 443  
 Db 362 GTATACACATTCGCGCGGCGCATGTACACAGCTTGACATCTGTGCAACCA 421  
 Qy 444 CCCCTTCCCGCCATGTGCGCAGGCTGTCTCATCTTCAAGGGATGTCTTCACTCGTGT 503  
 Db 422 CCCCTTCCCGCCATGTGCGCAGGCTGTCTCATCTTCAAGGGATGTCTTCACTCGTGT 481  
 Qy 504 CAGCTGATATACCGGAGTCTTGTGATGAGTCAAGCGCTTCAAGAGCAGATCTTGA 563  
 Db 482 CAGCTGATATACCGGAGTCTTGTGATGAGTCAAGCGCTTCAAGAGCAGATCTTGA 541  
 Qy 564 GGTCAAGTCTGCTGTGAAGAACAGACCAAGAGCGCGGAGCTGCCATGTGATCTG 623  
 Db 542 GGTCAAGTCTGCTGTGAAGAACAGACCAAGAGCGCGGAGCTGCCATGTGATCTG 601  
 Qy 624 TGGCAACAAGAACAGACCGCGGAG 648  
 Db 602 TGGCAACAAGAACAGACCGCGGAG 626

RESULT 14  
 BE389944 623 bp mRNA linear EST 21-JUL-2000  
 LOCUS 601282582F1 NIH\_MGC\_44 Homo sapiens cDNA clone IMAGE:3604196 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE389944  
 VERSION BE389944.1 GI:9335309  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 623)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE Unpublished (1999)  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov

Plate: L1CM250 row: d column: 21  
High quality sequence stop: 614.  
Location/Qualifiers

## FEATURES

source

1..623  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3604196"  
/tissue\_type="endometrium, adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NH MGC 44"  
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Query Match 18.6%; Score 574.4; DB 10; Length 623;  
Best Local Similarity 98.5%; Pred. No. 6.6e-55;  
Matches 611; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

26 GCGGCGCCAGCCGCGCGTCCGAGCAGCGGAGGATCCCGCGCAGTACCCGCG 85  
4 GTGCGGCCAGCCCGCGCGT-CCGAGCAGCGGAGGATCCCGCGCAGTACCCGCG 62  
86 AGCCACACA-GACTCTGGAGAGCTCGCGCTGGAGCAGCAGGAGCTCCCGAGCTC 144  
63 AGCCACACATGACTCTGGAGAGCTCGCGCTGGAGCAGCAGGAGCTCCCGAGCTC 122  
145 CCGGCGCTCCGAGCAGCTCTGAGCCGTGGCAAGAGCCCGCGCGCATTCGACGCC 204  
123 CCGGCGCTCCGAGCAGCTCTGAGCCGTGGCAAGAGCCCGCGCGCATTCGACGCC 182  
205 CGAGCAGATGATGATCTTGTCCAGCGGAACTGACGCTGATGCTGCGCCGCAAAAC 264  
183 CGAGCAGATGATGATCTTGTCCAGCGGAACTGACGCTGATGCTGCGCCGCAAAAC 242  
265 TCATACCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 324  
243 TCATACCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302  
325 TTCCCATGAGCGCGCTTGGAGAGCAGTACACACCCACCATGAGAGCTTCCACCGTAG 384  
303 TTCCCATGAGCGCGCTTGGAGAGCAGTACACACCCACCATGAGAGCTTCCACCGTAG 362  
385 GTATACACATCCGCGCGCAGTACACACCATGAGAGCTTCCACCGTAG 444  
363 GTATACACATCCGCGCGCAGTACACACCATGAGAGCTTCCACCGTAG 422  
445 CCGTTCGCGCGCATGCGGAGCTGTCATCTCAAGAGGATGCTTCACTCTGCTGCTTC 504  
423 CCGTTCGCGCGCATGCGGAGCTGTCATCTCAAGAGGATGCTTCACTCTGCTGCTTC 482  
505 AGCCGAGTAAACCGGAGAGCTTCAAGAGGATGCTTCAAGAGGATGCTTCAAGAGG 564  
483 AGCCGAGTAAACCGGAGAGCTTCAAGAGGATGCTTCAAGAGGATGCTTCAAGAGG 542  
565 GTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624  
543 GTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601  
625 GGCAACCAAGACGACCGCG 644  
602 TGGCACAAGACGACCGCG 621

RESULT 15  
LOCUS BM796105 574 bp mRNA linear EST 05-MAR-2002  
DEFINITION K-BST0078737 S22SNUI6n1 Homo sapiens cDNA clone S22SNUI6n1-53-A09

5', mRNA sequence.  
BM796105  
BM796105.1 GI:19144337  
EST.

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Homo sapiens

## REFERENCE

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

## AUTHORS

1 (bases 1 to 574)

## TITLE

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

## JOURNAL

Oh, K.U., Cheong, J.E., Sohn, H.Y., Kim, J.W., Park, H.S., Kim, S. and

## COMMENT

Kim, Y.S.

## JOURNAL

21C Frontier Korean EST Project 2001

## CONTACT

Unpublished (2002)

## COMMENT

Contact: Kim YS

## COMMENT

Genome Research Center

## COMMENT

Korea Research Institute of Bioscience &amp; Biotechnology

## COMMENT

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

## COMMENT

Tel: +82-42-860-4470

## COMMENT

Fax: +82-42-860-4409

## COMMENT

Email: yongsung@mail.krdb.re.kr

## COMMENT

Plate: 53 row: A column: 09

## COMMENT

High quality sequence stop: 574.

## FEATURES

source

1..574  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S22SNUI6n1-53-A09"  
/sex="F"  
/tissue\_type="Ascites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-16"  
/lab\_host="DH10B"  
/clone\_1ib="S22SNUI6n1"  
/note="Organ: Stomach; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The S22SNUI6 library was contributed by the Soares laboratory and it was constructed as described by Ronald, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cells of SNU-16 culture. SNU-16 cell was obtained from Korean Cell Line Bank (KCLB). SNU-16 was established from ascitic fluids of Korean patients by Park J.G. et al. (1990), Cancer Res 50: 2773-2780."

## ORIGIN

Query Match 18.5%; Score 569.2; DB 12; Length 574;  
Best Local Similarity 99.5%; Pred. No. 2.6e-54;  
Matches 571; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1743 GGAATCATTTGATGCTGTAAGAGCTGAGTACCTCCCTGAGCCAGGCGCTCACCGATT 1802  
1 GGAATCATTTGATGCTGTAAGAGCTGAGTACCTCCCTGAGCCAGGCGCTCACCGATT 60  
1803 CAGATCCAGGCGCTCCAGCCGCGAGCGCTTCTCTGCTGCCAAGAGGATTTCCGCTG 1862  
61 CAGATCCAGGCGCTCCAGCCGCGAGCGCTTCTCTGCTGCCAAGAGGATTTCCGCTG 120  
1863 GCGTCTTGGCAGCTGACATTTGACCTCGCGCATTTGAGCTCCAGAGGTTTACAGAAATTGC 1922  
121 GCGTCTTGGCAGCTGACATTTGACCTCGCGCATTTGAGCTCCAGAGGTTTACAGAAATTGC 180  
1923 ACAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1982  
181 ACAAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
241 TACACCGAGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
2043 ATTCAAGGTACAGCTCCAGCTTGAAGACAGAGGCTGCCCTTATCTGCTGCTTCTCAAGCA 2102  
301 ATTCAAGGTACAGCTCCAGCTTGAAGACAGAGGCTGCCCTTATCTGCTGCTTCTCAAGCA 360

```

QY 2103 TTATTAAGCACTACTGAGGTGCTGGTTCACTGTCCTAGAAACCAAGAGGGTCCC 2162
Db 361 TTTATTAAGCACTACTGAGGTGCTGGTTCACTGTCCTAGAAACCAAGAGGGTCCC 420
QY 2163 AGTCCTGAGCTCTGAGCCGCGCCCTGCTGCGCCACCACTTCTGCACACAGCGGTGGGA 2222
Db 421 AGTCCTGAGCTCTGAGCCGCGCCCTGCTGCGCCACCACTTCTGCACACAGCGGTGGGA 480
QY 2223 GCGCGGAGAGAGCACTGGGACCCAGAACTGAGCTGGGAGGATCCGACAGAAAAGCTC 2282
Db 481 GCGCGGAGAGATCACTGGGACCCAGAACTGAGCTGGGAGGATCCGACAGAAAAGCTC 540
QY 2283 AGGCGGGGTCTTCTCCTTGTGCGCGGATGGGC 2316
Db 541 AGGCGGGGTCTTCTCCTTGTGCGCGGATGGGC 574

```

Search completed: June 20, 2004, 08:39:15  
 Job time : 7584 secs

THIS PAGE BLANK (USPTO)